



## On parabolic final value problems

Christensen, Ann-Eva

*Publication date:*  
2013

*Document Version*  
Early version, also known as pre-print

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*  
Christensen, A-E. (2013). *On parabolic final value problems*. Department of Mathematical Sciences, Aalborg University. Ph.D. Report Series

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### Take down policy

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.

PhD Report Series - No. 23 - 2013

---

**On Parabolic Final Value Problems**

Ann-Eva Christensen

---

AALBORG UNIVERSITY  
Department of Mathematical Sciences

---

## On Parabolic Final Value Problems

Ann-Eva Christensen

---

*Thesis submitted:* July 5, 2013

*Thesis defended:* October 18, 2013

*PhD degree conferred:* December 19, 2013

*PhD supervisor:* Associate Professor Jon E. Johnsen  
Aalborg University

*PhD committee:* Professor Joachim Toft  
Linnæus University

Professor Michael Pedersen  
Technical University of Denmark

Professor Arne Jensen  
Aalborg University

DEPARTMENT OF MATHEMATICAL SCIENCES

Fredrik Bajers Vej 7G  
9220 Aalborg, Denmark  
<http://www.math.aau.dk>

# Preface

This thesis is the result of the work in relation to my Ph.D.-project, supervised by Jon Johnsen and carried out at Aalborg University, Department of Mathematical Sciences, in the period 15/8-2008 to 1/7-2013.

I wish to thank Jon, to whom I'm greatly indebted, for his many inputs and suggestions as well as providing guidance in many other aspects regarding the Ph.D.

Ann-Eva Christensen  
Aalborg University, July 5, 2013

# Summary

The topic of this thesis is parabolic problems. In particular, we focus on final value problems and their well-posedness. Since these in general are thought of as ill-posed, there are not many contributions addressing this subject in the literature.

Chapter 1 contains a short introduction via a motivating example.

In Chapter 2, we start by introducing the spaces and operators to consider. Then we recall two well-known theories used to study the initial value problem, specifically a functional analysis approach, and an approach via semigroup theory. We end up by showing that a combination of the two turns out to be very useful, since it provides weak assumptions under which a mild solution from semigroup theory is in fact a solution in the sense of distributions.

In Chapter 3, we use the results from Chapter 2 to study the final value problem. The main result is the existence of a bijective correspondence between initial and the terminal data. This correspondence yields a well-posedness result for the final value parabolic problem.

In Chapter 4, we study the special case of the heat equation. In particular, we see that the final value heat equation with homogeneous boundary data is well-posed as a consequence of the previous results, and we prove a similar result for the final value heat equation with inhomogeneous boundary data.

We conclude by making some short remarks in Chapter 5 regarding some applications beyond the scope of, but based on, the well-posedness results proven earlier.

Various notions and notation used throughout the text without further explanation can be found in Appendix A.

# Danish summary

## Titel: Om parabolske slutværdiproblemer

Emnet for denne afhandling er parabolske problemer. I særdeleshed fokuserer vi på slutværdiproblemer og deres velstillethed. Da disse generelt ikke opfattes som velstillede, er der ikke mange bidrag at finde herom i litteraturen.

Kapitel 1 indeholder en kort introduktion via et motiverende eksempel.

I Kapitel 2 lægger vi ud med at introducere de rum og operatorer vi beskæftiger os med. Dernæst genkalder vi os to tilgange til begyndelsesværdiproblemet; én via funktionalanalytise og én via semigruppeteori. Vi slutter med at vise, at en kombination af disse to viser sig meget nyttig, idet den giver svage forudsætninger, under hvilke den milde løsning fra semigruppeteorien faktisk er en løsning i distributions forstand.

I Kapitel 3 benytter vi resultaterne fra Kapitel 2 til at studere slutværdiproblemet. Hovedresultatet her er eksistensen af en bijektiv korrespondence imellem begyndelses- og slutdata. Denne sammenhæng giver et velstillethedsresultat for de parabolske slutværdiproblemer. I den forbindelse giver vi en karakterisering af mængden af tilladelige data, og denne vises at udgøre et Hilbertrum udstyret med en passende norm.

I Kapitel 4 studerer vi varmeledningsligningen som special tilfælde. Specielt ser vi at slutværdiproblemet for den homogene varmeledningsligning er velstillet som direkte konsekvens af teorien i tilfældet med Lax-Milgram operatorer. Og desuden ser vi, at lignende konklusioner gælder i det inhomogene tilfælde. Hertil benytter vi lineariteten af systemet, og et passende valg af højreinvert til sporoperatoren, til at omskrive problemet til et ækvivalent slutværdiproblem med homogene randdata.

Afslutningsvist kommer vi med nogle korte kommentarer i Kapitel 5 vedrørende nogle anvendelser, der falder uden for, men som er baseret på, velstillethedsresultaterne, der tidligere er vist. Vi giver eksempelvis et kort bevis for at varmeledningsligningen er approksimativt kontrollerbar, når kontrolfunktionen virker på hele det indre af området.

Forskellige begreber og notation, der bliver benyttet igennem teksten uden yderligere forklaring er at finde i Appendiks A.



# Contents

Preface . . . . .	i
Summary . . . . .	ii
Danish Summary . . . . .	iii
<b>1. Introduction</b>	<b>1</b>
<b>2. Review of the initial value problem</b>	<b>2</b>
2.1. Spaces and operators . . . . .	2
2.2. Existence and uniqueness of a solution . . . . .	8
2.3. Semigroup theory . . . . .	14
2.4. A representation of the solution . . . . .	22
2.5. Notes . . . . .	23
<b>3. Well-posedness of the final value evolution problem</b>	<b>25</b>
3.1. A bijection between initial and terminal data . . . . .	25
3.2. Well-posedness of the final value problem . . . . .	31
3.3. Notes . . . . .	34
<b>4. Applications to the heat equation</b>	<b>36</b>
4.1. Homogeneous boundary data . . . . .	36
4.2. Inhomogeneous boundary data . . . . .	38
4.3. Notes . . . . .	45
<b>5. A few examples</b>	<b>46</b>
5.1. A study of the two semihomogeneous problems . . . . .	46
5.2. The heat equation with inner control . . . . .	48
5.3. Notes . . . . .	49
<b>Conclusion and outlook</b>	<b>51</b>
<b>A. Notions and notation</b>	<b>52</b>
<b>Bibliography</b>	<b>56</b>





# 1. Introduction

There is a wide range of applications to the field of solving parabolic evolution problems backwards in time. For example, say a nuclear power plant have been down for a couple of days due to a power failure. With a measurement of the present core temperature available, but with no available measurements during the power interruption, it would be essential to deduce information about the past core temperatures to decide whether a melt down has occurred or not.

However, this is not an easy problem, as the heat equation has a regularizing effect, and the problem is in general considered to be ill-posed. But some conclusions can be made as we shall see, although we need to impose severe conditions on the final data, which may not be applicable for practical purposes.

We shall however discuss the well-posedness of the final value parabolic problem from a theoretical view point in the present thesis.

## 2. Review of the initial value problem

This chapter concerns the following initial value problem

$$\begin{cases} u' + \mathcal{A}u = f \\ u(0) = u_0, \end{cases} \quad (2.0.1)$$

the solvability and well-posedness of which is well known and well studied. In the present chapter, we recall two approaches to study this problem; namely, we present an existence and uniqueness theorem in Section 2.2 and give a brief reminder on semigroup theory in Section 2.3.

These two approaches are then combined in Section 2.4 to yield an explicit representation of the solution in the general setting described in Section 2.1 below.

### 2.1. Spaces and operators

We let  $V$  and  $H$  be Hilbert spaces, such that  $V$  is a subspace of  $H$ , and

$$V \subseteq H \equiv H^* \subseteq V^*, \quad (2.1.1)$$

densely with continuous injections so that for  $v \in V$ ,

$$\|v\|_{V^*} \leq C_1 \|v\|_H \leq C_2 \|v\|_V. \quad (2.1.2)$$

Concerning  $V$  and  $H$ , we then say that  $V \subseteq H$  algebraically, topologically and densely. Furthermore, we assume that  $V$  is separable. We emphasize that special constants denoted  $C_j$ ,  $j \in \mathbb{N}$ , refer to specific constants throughout the text, the value of which is defined in the equation where the constant is introduced. Whereas constants denoted by  $c$  vary with the context.

For simplicity, we shall hereafter denote by  $(\cdot | \cdot)$  the inner product on  $H$ , and  $\|\cdot\|$  the corresponding norm. The scalar product on  $V$  and  $V^*$ , written  $\langle \cdot, \cdot \rangle_{V^*, V}$ , is often abbreviated by  $\langle \cdot, \cdot \rangle$  in the sequel. Note that it is sesquilinear in that we let  $V^*$  consist of the conjugated linear bounded functionals on  $V$  (cf. Appendix A). This convention is useful in computations later on. Therefore, the operator norm of this functional gives a norm on  $V^*$ , so

$$|\langle u, v \rangle| \leq \|u\|_{V^*} \|v\|_V. \quad (2.1.3)$$

---

Moreover, when  $u \in H$ , then  $\langle u, v \rangle = (u | v)$  for all  $v \in V$  by (2.1.1).

Later, we need to be able to differentiate the scalar product on  $V$  and  $V^*$ , and include the following lemmas for completeness.

**Lemma 2.1.1.** *For  $f, g \in C^1([0, T]; H)$ ,*

$$\frac{d}{dt} (f(t) | g(t)) = (f'(t) | g(t)) + (f(t) | g'(t)). \quad (2.1.4)$$

*Proof.* Let  $f$  and  $g$  be differentiable with  $f(t+h) - f(t) = ha + o(h)$ ,  $a \in V$  and  $g(t+h) - g(t) = hb + o(h)$ ,  $b \in V$ . By sesquilinearity we have that for every fixed  $t$ , and  $t+h \in [0, T]$ ,

$$\begin{aligned} (f(t+h) | g(t+h)) - (f(t) | g(t)) &= (f(t+h) - f(t) | g(t+h) - g(t)) \\ &\quad + (f(t) | g(t+h) - g(t)) \\ &\quad + (f(t+h) - f(t) | g(t)) \\ &= (f(t) | hb) + (ha | g(t)) + o(h) \\ &= h \left( (f(t) | g'(t)) + (f'(t) | g(t)) \right) + o(h), \end{aligned} \quad (2.1.5)$$

since e.g.  $(o(h) | o(h)) = h^2 (o(1) | o(1)) = o(h)$ . This shows (2.1.4).  $\square$

The above lemma is a classic result, and next we present a well-known generalization, see e.g. [Tem01, Lemma 3.1.2] in the case where  $u = v$ .

**Lemma 2.1.2.** *For  $u, v \in L_2(0, T; V) \cap H^1(0, T; V^*)$ , the  $L_1$ -function  $t \mapsto (u(t) | v(t))$  has distributional derivative*

$$\frac{d}{dt} (u | v) = \langle u', v \rangle + \langle v', u \rangle. \quad (2.1.6)$$

*Furthermore,  $u$  and  $v$  have continuous representatives on  $[0, T]$  with values in  $H$ .*

*Proof.* The proof is by regularization. Let  $u, v \in L_2(0, T; V)$  with distributional derivatives  $u', v' \in L_2(0, T; V^*)$ . Denote by  $\tilde{u}$  the extension of  $u$  by zero outside  $[0, T]$  and let  $\tilde{u}_m := \int_{\mathbb{R}} h_m(s) \tilde{u}(t-s) ds$ , where  $h_m$  is a cut-off function on  $\mathbb{R}$  with  $h_m(t) := mh(mt)$ ,  $\int h dt = 1$  and  $\text{supp } h \subseteq \overline{B(0, 1)}$ . Then,  $\tilde{u}_m$  is  $C^\infty(\mathbb{R}; V)$ ,  $\tilde{u}_m \rightarrow \tilde{u}$  in  $L_2(\mathbb{R}; V)$ , and  $\tilde{u}'_m \rightarrow \tilde{u}'$  in  $L_2(\mathbb{R}; V^*)$  as  $m \rightarrow \infty$ . We denote by  $u_m$  the restriction of  $\tilde{u}_m$  to  $[0, T]$ . Then,  $u_m \rightarrow u$  in  $L_2(0, T; V)$ , and because of the jump at the endpoints of the interval,  $u'_m \rightarrow u'$  in  $L_{2, \text{loc}}(0, T; V^*)$ . Similarly,  $v$  gives rise to  $v_m$ .

---

Since  $(u | v)$  is continuous in both entries, it is measurable whenever  $u, v \in L_1(0, T; V)$ , and since  $\int_0^T |(u | v)| dt < \infty$ , it makes sense to take the distributional derivative. By Lemma 2.1.1,

$$\frac{d}{dt} (u_m | v_m) = \langle u'_m, v_m \rangle + \langle u_m, v'_m \rangle. \quad (2.1.7)$$

In particular this holds in the sense of scalar distributions.

Now,  $(u_m | v_m) \rightarrow (u | v)$  in  $L_1(0, T)$ , and since taking distributional derivatives is a continuous operation,  $\frac{d}{dt} (u_m | v_m) \rightarrow \frac{d}{dt} (u | v)$  in  $\mathcal{D}'(0, T)$ .

Moreover  $\langle u'_m, v_m \rangle \rightarrow \langle u', v \rangle$  and  $\langle v'_m, u_m \rangle \rightarrow \langle v', u \rangle$  in  $\mathcal{D}'(0, T)$ . Hence, it holds in  $\mathcal{D}'(0, T)$  that

$$\frac{d}{dt} \langle u, v \rangle = \langle u', v \rangle + \langle v', u \rangle. \quad (2.1.8)$$

This holds in particular for  $u \equiv v$ , so it follows from [Tem01, Lemma 3.1.2] that  $u$  has a continuous representation. Likewise  $v$  has a continuous representation.  $\square$

We also need to recall the classical construction of the Lax-Milgram operator we consider in the sequel and let  $a(\cdot, \cdot)$  be a bounded  $V$ -elliptic sesquilinear form on  $V$ , i.e. the following inequalities hold for  $u, v \in V$ :

$$|a(u, v)| \leq C_3 \|u\|_V \|v\|_V \quad (2.1.9)$$

$$\operatorname{Re} a(v, v) \geq C_4 \|v\|_V^2. \quad (2.1.10)$$

Then we exploit that there exist an isometric and bijective correspondence between bounded sesquilinear forms  $v(\cdot, \cdot)$  on  $V$  and bounded operators  $B$  on  $V$  given by  $v(u, v) = (Bu | v)_V$ . Here isometric is in the sense that the operator norm of  $B$  equals the following norm on the sesquilinear form  $|v| = \sup \{v(u, v) \mid \|u\|_V = 1 = \|v\|_V\}$ . Now we introduce the operator  $\mathcal{A}_0 \in \mathbb{B}(V)$  given by

$$a(u, v) = (\mathcal{A}_0 u | v)_V \quad \forall u, v \in V. \quad (2.1.11)$$

The form defined by

$$a^*(v, u) := (v | \mathcal{A}_0 u)_V \quad (2.1.12)$$

is bounded and sesquilinear, hence has an associated operator  $\mathcal{A}_0^* \in \mathbb{B}(V)$  defined by  $(\mathcal{A}_0^* v | u)_V = a^*(v, u)$ . Therefore,  $(\mathcal{A}_0 u | v)_V = \overline{a^*(v, u)} = (u | \mathcal{A}_0^* v)_V$ , so  $\mathcal{A}_0^*$  is the adjoint of  $\mathcal{A}_0$  on  $V$ , and we say that  $a^*(v, u) = \overline{a(u, v)}$  is the adjoint of the sesquilinear form  $a(\cdot, \cdot)$ . Note that (2.1.10) immediately implies that  $\mathcal{A}_0$  and  $\mathcal{A}_0^*$  are injective. Furthermore, it implies that the lower bounds of  $\mathcal{A}_0$  and  $\mathcal{A}_0^*$  are positive, so  $\mathcal{A}_0$  and  $\mathcal{A}_0^*$  are in fact bijective (cf. [Gru09, Theorem 12.9]).

---

The Riesz representation theorem states that, for all  $v^* \in V^*$ , there exists  $\tilde{v} \in V$ , and a bijective isometry  $J \in \mathbb{B}(V, V^*)$ , such that  $v^* = J\tilde{v}$ , and  $\langle J\tilde{v}, v \rangle = (\tilde{v} | v)_V$  for all  $v \in V$ .

Now we define the operator  $\mathcal{A} \in \mathbb{B}(V, V^*)$  by  $\mathcal{A} := J \circ \mathcal{A}_0$ , and note that by (2.1.11)

$$\langle \mathcal{A}u, v \rangle = a(u, v), \quad \forall u, v \in V. \quad (2.1.13)$$

Similarly,  $\mathcal{A}' := J \circ \mathcal{A}_0^*$ . Then, as compositions of bijective maps,  $\mathcal{A}$  and  $\mathcal{A}'$  are bijective.

By restricting  $\mathcal{A}$  to an operator in  $H$ , we define the Lax-Milgram operator  $A$ , which possibly is an unbounded operator. More precisely, we define  $A$  by

$$\begin{aligned} D(A) &= \mathcal{A}^{-1}(H), \\ Av &= \mathcal{A}v, \quad \text{for } v \in D(A), \end{aligned} \quad (2.1.14)$$

with  $\mathcal{A}^{-1}(H)$  denoting the pre-image of  $H$  under  $\mathcal{A}$ . Similarly, it follows from the proof of [Gru09, Theorem 12.18], that the adjoint  $A^*$  in  $H$ , equals the restriction of  $\mathcal{A}'$  to an operator in  $H$ .

In the sequel  $A$  denotes the Lax-Milgram operator on  $H$  defined above, and  $\mathcal{A} : V \rightarrow V^*$  the bounded extension of  $A$ . We are interested in the following evolution problem for a vector function  $\mathbb{R}_+ \rightarrow H$ , denoted by  $u(t)$ :

$$\begin{cases} u'(t) + \mathcal{A}u(t) = f(t) \\ u(0) = u_0 \end{cases} \quad (2.1.15)$$

Often conclusions can be strengthened, or proved in a simplified way, if we have an orthonormal basis on  $H$  consisting of eigenvectors of  $A$ . We conclude this section with some additional remarks for  $A$  selfadjoint. In particular, we shall see that, under the additional assumptions that  $V \hookrightarrow H$  is compact, we in fact then have orthonormal bases of (scaled) eigenvectors not only in  $H$ , but also in  $V$  and  $V^*$ .

**Remark 2.1.3.** *When  $A$  is selfadjoint, and  $V \hookrightarrow H$  is a compact embedding, then  $H$  has an orthonormal basis of eigenvectors for  $A$ . Let us recall this classic fact.*

*Since  $a(\cdot, \cdot)$  is  $V$ -elliptic,  $0 \in \rho(A)$ , so that  $A : D(A) \rightarrow H$  is a bijection with  $A^{-1} \in \mathbb{B}(H)$ . Moreover,  $A^{-1}$  is bounded as a mapping  $H \rightarrow D(A)$  with respect to the graph topology, i.e. the topology defined by the graph norm*

$$\|v\|_{D(A)}^2 = \|v\|^2 + \|Av\|^2, \quad \text{for all } v \in D(A), \quad (2.1.16)$$

*for obviously*

$$\|A^{-1}f\|_{D(A)}^2 = \|A^{-1}f\|^2 + \|AA^{-1}f\|^2 \leq (c+1)\|f\|^2, \quad \text{for all } f \in H. \quad (2.1.17)$$

---

Normed this way,  $D(A)$  is a Banach space, as  $\mathcal{A}$  is closed, and the injection  $D(A) \hookrightarrow V$  is bounded, which is seen, from  $V$ -ellipticity, as follows:

$$C_4 \|v\|_V^2 \leq \operatorname{Re} a(v, v) \leq |(Av \mid v)| \leq \frac{1}{2} (\|Av\|^2 + \|v\|^2) = \frac{1}{2} \|v\|_{D(A)}^2. \quad (2.1.18)$$

Hence, as the composition of a bounded map and a compact map,  $A^{-1} : H \rightarrow D(A) \hookrightarrow V \hookrightarrow H$  is compact whenever  $V \hookrightarrow H$  is so. Since  $A$  is selfadjoint, closed, and densely defined with dense range in  $H$ ,  $A^{-1}$  is selfadjoint (cf. [Gru09, Theorem 12.7]). The spectral theorem for compact selfadjoint operators implies that  $H$  has an orthonormal basis consisting of eigenvectors of  $A^{-1}$ , and that  $\sigma(A^{-1}) \subseteq \{0\} \cup \{\mu_j \mid A^{-1}e_j = \mu_j e_j, e_j \neq 0\}$ , where the eigenvalues  $\mu_j$  of  $A^{-1}$  can be ordered such that

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_j \geq 0, \quad \text{with } \mu_j \searrow 0, j \rightarrow \infty. \quad (2.1.19)$$

Note that  $0 \notin \sigma_{\text{point}}(A^{-1})$  but that  $0 \in \sigma_{\text{cont}}(A^{-1})$  could occur, since  $A$  may be unbounded. We shall show that this is not the case since  $A$  is  $V$ -elliptic. Now,

$$\mu \text{ eigenvalue of } A^{-1} \Leftrightarrow \frac{1}{\mu_j} \text{ eigenvalue of } A, \quad (2.1.20)$$

and the eigenvectors are the same. Consequently,  $H$  has an orthonormal basis consisting of eigenvectors of  $A$ . Moreover,  $\sigma(A) = \{\lambda_j\} := \{\mu_j^{-1}\}$ . Indeed  $\sigma_{\text{res}}(A)$  is empty since  $A$  is selfadjoint. Also,  $\sigma_{\text{point}}(A) = \{\lambda_j\}$  and  $0 \notin \sigma(A)$ . When  $\nu \neq 0$ ,  $\nu \neq \frac{1}{\lambda_j}$ , then  $\nu \notin \sigma_{\text{cont}}(A)$  since the operator

$$(A - \nu I) = (\nu^{-1} - A^{-1})\nu A$$

has bounded inverse, as  $\frac{1}{\nu} \notin \sigma(A^{-1})$ . Therefore,  $\sigma(A) = \sigma_{\text{point}}(A) = \{\lambda_j\}$ .  $\circ$

Proceeding from Remark 2.1.3, we note that  $a(\cdot, \cdot)$  is an inner product on  $V$ , since it is symmetric and  $V$ -elliptic. Hence, it induces a norm on  $V$

$$\|v\|_V^2 := a(v, v), \quad (2.1.21)$$

which by the boundedness and  $V$ -ellipticity of  $a(\cdot, \cdot)$  is equivalent to the usual norm  $\|v\|_V$ . This is analysed in

**Fact 1.** For all  $v \in V$ ,  $\sum_{j=1}^{\infty} (v \mid e_j) e_j$  converges not only in  $H$ , but in  $V$  too. Furthermore,  $\{\lambda_j^{-1/2} e_j\}$  is an orthonormal basis for  $V$  and  $\|v\|_V^2 = \sum_{j=1}^{\infty} \lambda_j | (v \mid e_j) |^2$ .

*Proof.* Note that

$$a(e_j, e_k) = \langle \mathcal{A}e_j, e_k \rangle = \lambda_j (e_j \mid e_k), \quad (2.1.22)$$

---

so that  $\{\lambda_j^{-1/2}e_j\}$  is an orthonormal sequence in  $V$ . It is in fact a basis, since  $\{e_j\}$  is a basis on  $H$  (see e.g. [Ped00, Chapter 5]). Since  $a(\cdot, \cdot)$  is symmetric and  $V$ -elliptic,  $\lambda_j > 0$ . Hence, for all  $v \in V$ ,

$$v = \sum_{j=1}^{\infty} a(v, \lambda_j^{-1/2}e_j) \lambda_j^{-1/2}e_j = \sum_{j=1}^{\infty} \overline{a(e_j, v)} \lambda_j^{-1}e_j = \sum_{j=1}^{\infty} (v | e_j) e_j, \quad (2.1.23)$$

so that the rightmost side converges in  $V$ .

Moreover,

$$v = \sum_{j=1}^{\infty} \lambda_j^{1/2} (v | e_j) \lambda_j^{-1/2}e_j \quad (2.1.24)$$

is an orthogonal expansion of  $v$ , which implies

$$\|v\|_V^2 = \sum_{j=1}^{\infty} \lambda_j | (v | e_j) |^2. \quad (2.1.25)$$

□

Since  $\mathcal{A}$  is a bijection, we note that by the symmetry and  $V$ -ellipticity of  $a(\cdot, \cdot)$ ,

$$(w_1 | w_2)_{V^*} := a(\mathcal{A}^{-1}w_1, \mathcal{A}^{-1}w_2) = \langle w_1, \mathcal{A}^{-1}w_2 \rangle \quad (2.1.26)$$

is an inner product on  $V^*$ , and hence induces a norm

$$\|w\|_{V^*}^2 := a(\mathcal{A}^{-1}w, \mathcal{A}^{-1}w), \quad (2.1.27)$$

which can be shown to be equivalent to the usual norm  $\|w\|_{V^*}^2$ , by the boundedness of  $\mathcal{A}$  and  $\mathcal{A}^{-1}$ . Indeed,

$$\|w\|_{V^*}^2 = |\langle w, \mathcal{A}^{-1}w \rangle| \leq \|\mathcal{A}^{-1}\|_{\mathbb{B}(V^*, V)} \|w\|_{V^*}^2, \quad (2.1.28)$$

and by Fact 1

$$\|w\|_{V^*}^2 \leq \|\mathcal{A}\|_{\mathbb{B}(V, V^*)}^2 \|\mathcal{A}^{-1}w\|_V^2 \leq c' \|\mathcal{A}^{-1}w\|_V^2 = c' \|w\|_{V^*}^2. \quad (2.1.29)$$

Similar to Fact 1, we have

**Fact 2.** For all  $w \in V^*$ ,  $\sum_{j=1}^{\infty} \langle w, e_j \rangle e_j$  converges not only in  $H$ , but in  $V^*$  too. Furthermore,  $\{\lambda_j^{1/2}e_j\}$  is an orthonormal basis of  $V^*$  and  $\|w\|_{V^*}^2 = \sum_{j=1}^{\infty} \lambda_j^{-1} |\langle w, e_j \rangle|^2$ .



---

*Proof.* Since

$$(e_j \mid e_k)_{V^*} = \langle e_j, \mathcal{A}^{-1} e_k \rangle = \lambda_k^{-1} (e_j \mid e_k), \quad (2.1.30)$$

the sequence  $\{\lambda_j^{1/2} e_j\}$  is orthonormal in  $V^*$ . It is in fact a basis, as  $w \in V^*$  with  $0 = \langle e_j, \mathcal{A}^{-1} w \rangle = (e_j \mid \mathcal{A}^{-1} w)$  for all  $j$ , implies that  $w = 0$  by the injectivity of  $\mathcal{A}^{-1}$ .

Now, all  $w \in V^*$  has an orthogonal expansion

$$w = \sum_{j=1}^{\infty} \langle w, \mathcal{A}^{-1} \lambda_j^{1/2} e_j \rangle \lambda_j^{1/2} e_j = \sum_{j=1}^{\infty} \langle w, \mathcal{A}^{-1} e_j \rangle \lambda_j e_j = \sum_{j=1}^{\infty} \langle w, e_j \rangle e_j, \quad (2.1.31)$$

so that the rightmost side is convergent in  $V^*$ . Since

$$w = \sum_{j=1}^{\infty} \langle w, e_j \rangle e_j = \sum_{j=1}^{\infty} \lambda_j^{-1/2} \langle w, e_j \rangle \lambda_j^{1/2} e_j, \quad (2.1.32)$$

we have that

$$\|w\|_{V^*}^2 = \sum_{j=1}^{\infty} \lambda_j |\langle w, e_j \rangle|^2. \quad (2.1.33)$$

□

We return to this setting later on, since it is sufficient for many applications. In particular, we consider this setting when studying the heat equation in Chapter 4.

## 2.2. Existence and uniqueness of a solution

Having set the scene, we present a result on the existence and uniqueness of a solution to the following problem

$$\begin{cases} u' + \mathcal{A}u = f & \text{in } \mathcal{D}'(0, T; V^*) \\ u(0) = u_0 & \text{in } H. \end{cases} \quad (2.2.1)$$

The result is inspired by [Tem01, Chapter 3], where the proof is presented in the case where  $a(u, v) = (u \mid v)_V$  and for specific spaces with relevance for the Navier-Stokes equations.

We state and prove the theorem in the general setting introduced in Section 2.1 following the same lines as the proof in [Tem01] and emphasize, that  $A$  need not be selfadjoint.

---

**Theorem 2.2.1.** *Let  $V$  be a separable Hilbert space with  $V \subseteq H$  algebraically, topologically and densely, cf. (2.1.1) and (2.1.2). Let  $\mathcal{A}$  be the operator introduced in (2.1.13). Assume that  $u_0 \in H$  and  $f \in L_2(0, T; V^*)$ . Then there exist a uniquely determined function*

$$u \in L_2(0, T; V) \cap C([0, T]; H) \cap H^1(0, T; V^*) \quad (2.2.2)$$

*satisfying (2.2.1).*

Before we give the proof, we recall the following important lemma (see [Tem01, Lemma 1.1, Chapter 3]) that e.g. allows one to integrate by parts in a general setting.

**Lemma 2.2.2.** *Let  $B$  be a given Banach space with dual  $B'$ , and let  $u, g \in L_1(a, b; B)$ . Then the following conditions are equivalent*

- *$u$  is a.e. equal to a primitive function of  $g$ ,*

$$u(t) = \xi + \int_a^t g(s) ds, \quad \xi \in B, \text{ for a.e. } t \in [a, b]. \quad (2.2.3)$$

- *For each test function  $\phi \in C_0^\infty ]a, b[$ ,*

$$\int_b^a u(t) \phi'(t) dt = - \int_b^a g(t) \phi(t) dt. \quad (2.2.4)$$

- *For each  $\eta \in B'$ ,*

$$\frac{d}{dt} \langle \eta, u \rangle_{B', B} = \langle \eta, g \rangle_{B', B} \quad (2.2.5)$$

*in the scalar distribution sense on  $]a, b[$ .*

*In the affirmative case, there is a continuous function  $v : [a, b] \rightarrow B$  such that  $u(t) = v(t)$  a.e. on  $[a, b]$ , and  $\xi = u(a)$ . Moreover,*

$$\sup_{a \leq t \leq b} \|u(t)\|_B \leq (b - a)^{-1} \|u\|_{L_1(a, b; B)} + \|g\|_{L_1(a, b; B)}. \quad (2.2.6)$$

This lemma is important to the proof of Theorem 2.2.1, since it implies that  $u$  solving (2.2.1) in  $\mathcal{D}'(0, T; V^*)$  is equivalent to  $u$  satisfying  $\partial_t \langle u, \eta \rangle + a(u, \eta) = \langle f, \eta \rangle$  for all  $\eta \in V$  in  $\mathcal{D}'(0, T)$ . Indeed, there is first of all the identification between  $V$  and the dual of  $V^*$ , and  $\langle \mathcal{A}u, \eta \rangle = a(u, \eta)$  by the definition of  $\mathcal{A}$ , cf. (2.1.13); whilst  $\langle u', \eta \rangle = \partial_t \langle u, \eta \rangle$  in view of (2.2.5) (which is exploited in the proof of Theorem 2.4.1 below).

Furthermore, the lemma will be a key ingredient in the proof of a representation formula for the solution to (2.2.1).

---

**Remark 2.2.3.** We remark that we will refer to this lemma even in the case where  $B = \mathbb{C}$ , instead of including a separate reference to the fundamental theorem of analysis for distributions.  $\circ$

**Remark 2.2.4.** The proof of Lemma 2.2.2 can be found in [Tem01], but let us make a short comment on the additional remark in (2.2.6). It can be seen by noting that the continuous function  $\|u(t)\|_B$  attains its minimum on the interval  $[a, b]$  for some  $t_0$ , so that by the Bochner inequality (see e.g. [Eva08]) and (2.2.3),

$$\sup_t \|u(t)\|_B \leq \|u(t_0)\|_B + \int_a^b \|g(t)\|_B dt. \quad (2.2.7)$$

Then an application of the mean value theorem for integrals yields (2.2.6).

If furthermore  $u, g \in L_2(a, b; B)$ , we have that  $(\|u(t_0)\|_B^2)^{1/2} \leq (b - a)^{-1/2} \|u\|_{L_2(0, T; B)}$  so that, by the Cauchy-Schwartz inequality

$$\sup_{a \leq t \leq b} \|u(t)\|_B \leq (b - a)^{-1/2} \|u\|_{L_2(a, b; B)} + (b - a)^{1/2} \|g\|_{L_2(a, b; B)} \leq c \|u\|_{H^1(a, b; B)}. \quad (2.2.8)$$

$\circ$

**Remark 2.2.5.** As a preparation for the proof of Theorem 2.2.1, let us note that the distributional derivative of a function with real values is itself real valued. Indeed, for  $u \in \mathcal{D}'(0, T)$  with real values on test functions  $\phi \in C_0^\infty([0, T]; \mathbb{R})$ , one has

$$\operatorname{Im} \langle u', \phi \rangle = -\operatorname{Im} \langle u, \phi' \rangle \equiv 0. \quad (2.2.9)$$

$\circ$

*Proof of Theorem 2.2.1.* The strategy is to use the Faedo-Galerkin method to construct an approximate solution, and then pass to the limit. Since  $V$  is separable, there exist a countable orthonormal basis for  $V$ , denoted  $(w_n)_{n \in \mathbb{N}}$ . Let  $w_1, \dots, w_m$  denote the first  $m$  basis vectors. We construct a  $u_m(t)$  that solves

$$\begin{cases} (u'_m | w_j) + a(u_m, w_j) = \langle f, w_j \rangle & \text{for } j = 1, \dots, m \\ u_m(0) = u_{0m}, \end{cases} \quad (2.2.10)$$

where  $u_{0m} \in V$  is chosen such that  $\|u_{0m} - u_0\| \rightarrow 0$  as  $m \rightarrow \infty$ , which is possible since  $V \subseteq H$  densely. Tentatively, we set

$$u_m = \sum_{i=1}^m g_{im}(t) w_i. \quad (2.2.11)$$

By inserting (2.2.11) in (2.2.10), and noting that the  $m \times m$  matrix  $[(w_i | w_j)]_{i,j \leq m}$  is invertible since the  $w_i$ 's are linearly independent, we see that (2.2.10) is equivalent to

$$g'_{im}(t) + \sum_{j=1}^m \alpha_{ij} g_{jm}(t) = \sum_{j=1}^m \beta_{ij} \langle f(t), w_j \rangle \quad \text{for } i = 1, \dots, m, \quad (2.2.12)$$

with  $[\beta]_{i,j \leq m} = [(w_i | w_j)]_{i,j \leq m}^{-1}$  and  $[\alpha]_{i,j \leq m} = [\beta]_{i,j \leq m} [a(w_i, w_j)]_{i,j \leq m}$ . This is a linear system with constant coefficients. Imposing the initial condition  $g_{im}(0) = u_{0m}^i$  for  $i = 1, \dots, m$ , where  $u_{0m}^i$  denotes the  $i$ 'th coordinate of  $u_{0m}$ , this system has a unique solution  $g_m(t) := (g_{1m}(t), \dots, g_{mm}(t)) \in C^0([0, T]; \mathbb{C}^m)$ . Hence,  $u_m$  defined in (2.2.11) is an approximate solution in the sense of (2.2.10). Since (2.2.10) holds for all  $j \in [1, m]$ , it follows by (2.2.11), and sesquilinearity that

$$(u'_m | u_m) + a(u_m, u_m) = \langle f(t), u_m \rangle. \quad (2.2.13)$$

By assumption,  $f \in L_2(0, T; V^*)$ , and  $w_j \in V$ , so  $\langle f(t), w_j \rangle \in L_2(0, T)$ . From (2.2.12)  $g_{im}, g'_{im} \in L_2(0, T)$ . Hence,  $u_m \in L_2(0, T; V)$ ,  $u'_m \in L_2(0, T; V) \subseteq L_2(0, T; V^*)$ , and by Lemma 2.1.2,

$$\frac{d}{dt} \|u_m\|^2 + 2a(u_m, u_m) = 2\langle f(t), u_m \rangle. \quad (2.2.14)$$

Now, by Remark 2.2.5,  $\frac{d}{dt} \|u_m\|^2$  has real values, so applying (2.1.3) to the right hand side of (2.2.14) and (2.1.10) to the left hand side, yields

$$\frac{d}{dt} \|u_m\|^2 + 2C_4 \|u_m\|_V^2 \leq 2\|f\|_{V^*} \|u_m\|_V. \quad (2.2.15)$$

Note that  $\|u_{0m}\| \leq \|u_{0m} - u_0\| + \|u_0\|$ , so that  $\|u_{0m}\| \leq 2\|u_0\|$  eventually. For any  $c > 0$

$$2\|f\|_{V^*} \|u_m\|_V \leq c^{-1} \|f\|_{V^*}^2 + c \|u_m\|_V^2. \quad (2.2.16)$$

Choosing  $c = 2C_4$ , applying (2.2.16) to the right hand side of (2.2.15), and integrating with respect to  $t$  by using Lemma 2.2.2, we get for  $s \in [0, T]$

$$\|u_m(s)\|^2 \leq \|u_m(0)\|^2 + (2C_4)^{-1} \|f\|_{L_2(0,s;V^*)}^2 \leq 4\|u_0\|^2 + (2C_4)^{-1} \|f\|_{L_2(0,T;V^*)}^2. \quad (2.2.17)$$

Taking the supremum over  $s \in [0, T]$  this shows that  $\|u_m(t)\|_{L_\infty(0,T;H)} \leq k$  for all  $t \in [0, T]$ . By the Banach-Alaouglu Theorem  $u_{m*} \rightarrow u$  in the weak\* topology on  $L_\infty(0, T; H)$ , where  $u \in L_\infty(0, T; H)$ , and  $\{u_{m*}\}$  denotes a subsequence of  $\{u_m\}$ , i.e.

$$\int_0^T (u_{m*} | v) dt \rightarrow \int_0^T (u | v) dt \quad \forall v \in L_1(0, T; H). \quad (2.2.18)$$

A similar computation, choosing  $c = C_4$  in (2.2.16), yields

$$\|u_m(T)\|^2 + C_4 \|u_m\|_{L_2(0,T;V)}^2 \leq 4\|u_0\|^2 + C_4^{-1} \|f\|_{L_2(0,T;V^*)}^2. \quad (2.2.19)$$

In particular,  $\|u_{m*}\|_{L_2(0,T;V)}^2 \leq k, \forall t \in [0, T]$ . The Banach-Alaouglu Theorem implies that  $u_{m'} \rightharpoonup u_*$  in the weak\* topology on  $L_2(0, T; V)$ , with  $u_* \in L_2(0, T; V)$  and  $\{u_{m'}\}$  denoting a subsequence of  $\{u_{m*}\}$ , i.e.

$$\int_0^T \langle v, u_{m'} \rangle dt \rightarrow \int_0^T \langle v, u_* \rangle dt \quad \forall v \in L_2(0, T; V^*). \quad (2.2.20)$$

By comparison with (2.2.18), we see that  $u$  and  $u_*$  coincides as functionals on  $v \in L_2(0, T; H)$ . Hence  $u = u_* \in L_2(0, T; V) \cap L_\infty(0, T; H)$ .

To show that  $u$  solves (2.2.1), we integrate (2.2.10) up against a smooth function  $\psi \in C^\infty([0, T])$  with  $\psi(T) = 0$ , by Lemma 2.2.2, to get

$$\int_0^T (u_m | w_j) \psi' dt + \int_0^T a(u_m, w_j) \psi dt = (u_{0m} | w_j) \psi(0) + \int_0^T \langle f, w_j \rangle \psi dt. \quad (2.2.21)$$

The first term on the left hand side, converges as  $m \rightarrow \infty$ , using (2.2.18) as  $\psi'(t)w_j \in L_1(0, T; H)$ . The second term on the left hand side converges using (2.2.20), by noting that  $\mathcal{A}'w_j\psi(t) \in L_2(0, T; V^*)$ , and

$$a(u_m, w_j) = (\mathcal{A}_0 u_m | w_j)_V = (u_m | \mathcal{A}_0^* w_j)_V = \overline{\langle \mathcal{A}' w_j, u_m \rangle}. \quad (2.2.22)$$

Also,  $u_{0m} \rightarrow u_0$  strongly, so in the limit,

$$\int_0^T (u | w_j) \psi' dt + \int_0^T a(u, w_j) \psi dt = (u_0 | w_j) \psi(0) + \int_0^T \langle f, w_j \rangle \psi dt. \quad (2.2.23)$$

Now we choose  $\psi$  to be a test function, use that  $v = \sum_{j=1}^\infty (v | w_j) w_j, \forall v \in V$ , and that  $(\cdot | \cdot), \langle \cdot, \cdot \rangle$ , and  $a(\cdot, \cdot)$  are continuous in both entries. Then, in the scalar distributional sense,

$$\frac{d}{dt} (u | v) + a(u, v) = \langle f, v \rangle \quad \forall v \in V. \quad (2.2.24)$$

As previously noted, since  $u, f, \mathcal{A}u \in L_2(0, T; V^*)$ , this is equivalent to  $u$  satisfying (2.2.1) by Lemma 2.2.2. Furthermore, the distributional derivative  $u' = f - \mathcal{A}u$  is in  $L_2(0, T; V^*)$ , so  $u \in H^1(0, T; V^*)$ .

It follows immediately from Lemma 2.1.2 that  $u = g$  for almost every  $t \in [0, T]$ , for a suitable  $g \in C([0, T]; H)$

---

Note that (2.2.24) holds in particular for  $v = w_i$ ,  $\forall i \geq 1$ , so integrating by parts, by Lemma 2.2.2, against  $\psi \in C^\infty(0, T)$  with  $\psi(T) = 0$ , and comparison with (2.2.23) gives that  $(u_0 - u(0) \mid w_i) \psi(0) = 0$ ,  $\forall i \geq 1$  since the other terms are the same. By choosing  $\psi(0) \neq 0$ , we see that  $(u_0 - u(0) \mid w_i) = 0$ ,  $\forall i$ , i.e.  $u_0 - u(0) \perp \text{Span}(w_n)_{n \in \mathbb{N}}$ . Since  $\text{Span}(w_n)_{n \in \mathbb{N}}$  is dense in  $V$ , it is dense in  $H$ . Hence,  $u$  satisfies the initial condition  $u(0) = u_0$ .

Uniqueness follows from Lemma 2.1.2, as we now show. Assuming  $u$  and  $v$  both solves (2.2.1),  $w := u - v$  satisfies  $w' + \mathcal{A}w = 0$ , and  $w(0) = 0$ . Hence,

$$\frac{d}{dt} \|w\|^2 + 2a(w, w) = 2\langle w', w \rangle + 2a(w, w) = 0 \quad (2.2.25)$$

Since  $\partial_t \|w\|^2 \in \mathbb{R}$  and  $a(w, w)$  is  $V$ -elliptic,  $\partial_t \|w\|^2 \leq 0$ , which implies that  $\|w(t)\|^2 \leq \|w(0)\|^2$  by integration. Hence  $u \equiv v$  for all  $t \in [0, T]$ .  $\square$

We say that a problem is well-posed if there exist a solution that is unique and depends continuously on the data. This notion of well-posedness was introduced by J. Hadamard (see e.g. [Had02]). Of course, the well-posedness of a problem depends heavily on the space in which the problem is considered. Next, we show that the solution from Theorem 2.2.1 depends continuously on the data, yielding that (2.2.1) is well-posed. In order to make this precise, let  $\|\cdot\|_X$  denote the norm on the Banach space

$$X := L_2(0, T; V) \cap C([0, T]; H) \cap H^1(0, T; V^*), \quad (2.2.26)$$

i.e.

$$\|u\|_X^2 = \|u\|_{L_2(0, T; V)}^2 + \sup_{0 \leq t \leq T} \|u(t)\|_H^2 + \|u\|_{H^1(0, T; V^*)}^2. \quad (2.2.27)$$

This notation will be used throughout the text. We have the following result.

**Corollary 2.2.6.** *The solution  $u$  to (2.2.1) depends continuously on  $f \in L_2(0, T; V^*)$ , and  $u_0 \in H$ , i.e.*

$$\|u\|_X^2 \leq c(\|u_0\|^2 + \|f\|_{L_2(0, T; V^*)}^2). \quad (2.2.28)$$

*Proof.* Let  $u$  be the solution from Theorem 2.2.1. Then  $u \in L_2(0, T; V)$ , while both  $u'$  and  $f$  is in  $L_2(0, T; V^*)$ , so  $\mathcal{A}u \in L_2(0, T; V^*)$  by (2.2.1), and

$$\langle u', u \rangle + \langle \mathcal{A}u, u \rangle = \langle f, u \rangle. \quad (2.2.29)$$

By Lemma 2.1.2, we have

$$\partial_t \|u\|^2 + 2a(u, u) = 2\langle f, u \rangle, \quad (2.2.30)$$

---

so that (cf. Remark 2.2.5)

$$\partial_t \|u\|^2 + 2 \operatorname{Re} a(u, u) = 2 \operatorname{Re} \langle f, u \rangle, \quad (2.2.31)$$

and by  $V$ -ellipticity,

$$\partial_t \|u\|^2 + 2C_4 \|u\|_V^2 \leq 2|\langle f, u \rangle| \leq C_4^{-1} \|f\|_{V^*}^2 + C_4 \|u\|_V^2. \quad (2.2.32)$$

Note that  $\partial_t \|u\|^2 \in L_1(0, T)$  by (2.2.30), so that Lemma 2.2.2 yields

$$\|u(t)\|^2 + C_4 \int_0^t \|u(s)\|_V^2 \, ds \leq \|u_0\|^2 + C_4^{-1} \|f\|_{L_2(0, T; V^*)}^2. \quad (2.2.33)$$

Hence,

$$\sup_{0 \leq t \leq T} \|u(t)\|^2 \leq \|u_0\|^2 + C_4^{-1} \|f\|_{L_2(0, T; V^*)}^2, \quad (2.2.34)$$

and

$$\|u\|_{L_2(0, T; V)}^2 \leq C_4^{-1} (\|u_0\|^2 + C_4^{-1} \|f\|_{L_2(0, T; V^*)}^2). \quad (2.2.35)$$

Since  $u$  solves (2.2.1),

$$\|\partial_t u(t)\|_{V^*}^2 \leq (\|f(t)\|_{V^*} + \|Au\|_{V^*})^2 \leq 2(\|f(t)\|_{V^*}^2 + \|Au\|_{V^*}^2). \quad (2.2.36)$$

This implies

$$\begin{aligned} \int_0^T \|\partial_t u(t)\|_{V^*}^2 \, dt &\leq 2 \int_0^T \|f(t)\|_{V^*}^2 \, dt + 2\|\mathcal{A}\|_{\mathbb{B}(V, V^*)}^2 \int_0^T \|u\|_V^2 \, dt \\ &\leq 2\|\mathcal{A}\|_{\mathbb{B}(V, V^*)}^2 C_4^{-1} \|u_0\|^2 + 2(1 + C_4^{-2} \|\mathcal{A}\|_{\mathbb{B}(V, V^*)}^2) \|f\|_{L_2(0, T; V^*)}^2, \end{aligned} \quad (2.2.37)$$

and shows (2.2.28).  $\square$

It is possible to draw further conclusions on the solution  $u$ . For this purpose, we include the next section.

## 2.3. Semigroup theory

To recall that when  $A$  is a Lax-Milgram operator on  $H$ , then  $-A$  generates an analytic semigroup in  $H$  (cf. Lemma 2.3.3 below), we include the definition for the sake of completeness (see e.g. [Paz83, Section 2.5]).

---

**Definition 2.3.1.** Let  $W$  be a Hilbert space, and  $S = \{0\} \cup \{\lambda \in \mathbb{C} \mid \phi_1 < \arg \lambda < \phi_2\}$ , with given angles  $\phi_1, \phi_2$  satisfying  $\phi_1 < 0 < \phi_2$ . Then  $T(\lambda)$ ,  $\lambda \in S$ , is an analytic semigroup if each  $T(\lambda)$  is a bounded linear operator on  $W$ ,  $\lambda \mapsto T(\lambda)$  is an analytic map  $S^\circ \rightarrow \mathbb{B}(W)$ , and we have the semigroup properties:

- (i)  $T(0) = I$ ,
- (ii)  $T(\lambda_1 + \lambda_2) = T(\lambda_1)T(\lambda_2)$ ,  $\forall \lambda_1, \lambda_2 \in S$ ,
- (iii)  $\lim_{\lambda \rightarrow 0} T(\lambda)x = x$ ,  $\lambda \in S, x \in W$ .

◇

It later becomes useful, that when  $T(t)$  is an analytic semigroup, or simply a strongly continuous semigroup of contractions, there exist constants  $\omega \geq 0$ , and  $M \geq 1$ , such that

$$\|T(t)\|_{\mathbb{B}(W)} \leq Me^{\omega t} \quad \text{for } 0 \leq t < \infty. \quad (2.3.1)$$

The following theorem is a well-known criterion for an operator to be a generator for an analytic semigroup. The proof can be found in [Paz83, Theorem 2.5.2] along with several equivalent criterions.

**Theorem 2.3.2.** If there exist  $0 < \theta < \frac{\pi}{2}$ , and  $M > 0$  such that

$$\rho(A) \supseteq \Sigma = \left\{ \lambda \in \mathbb{C} \mid |\arg \lambda| < \frac{\pi}{2} + \theta \right\} \cup \{0\}, \quad (2.3.2)$$

and

$$\|(\lambda I - A)^{-1}\|_{\mathbb{B}(W)} \leq \frac{M}{|\lambda|}, \quad \text{for } \lambda \in \Sigma, \lambda \neq 0, \quad (2.3.3)$$

then  $A$  generates an analytic semigroup  $T(\lambda)$  in the sector  $\{\lambda \in \mathbb{C} \mid |\arg \lambda| < \theta\}$ . Furthermore,  $T(t)$  is differentiable in  $\mathbb{B}(W)$ , when  $t > 0$ , with  $T'(t) = AT(t)$  satisfying the following bound

$$\|AT(t)\|_{\mathbb{B}(W)} \leq \frac{c}{t} \quad \text{for } t > 0. \quad (2.3.4)$$

Next, we show that Lax-Milgram operators are generators of analytic semigroups, up to a sign.

**Lemma 2.3.3.** Let  $A$  be a Lax-Milgram operator defined from the triple  $(V, H, a(\cdot, \cdot))$ , where  $a(\cdot, \cdot)$  is a bounded  $V$ -elliptic sesquilinear form, and  $V \subseteq H$  algebraically, topologically and densely. Then  $-A$  and  $(-A)^*$  generate analytic semigroups on  $H$ .



The proof is similar to the proof found of Lemma 3.6.1 in [Tan79] that covers a half plane, but since we need to extend it to a sector  $\{\lambda \in \mathbb{C} \mid |\arg \lambda| < \frac{\pi}{2} + \theta\}$ ,  $\theta \in ]0, \frac{\pi}{2}[$ , we present it here for completeness. First note, that the boundedness, and  $V$ -ellipticity of  $a(\cdot, \cdot)$ , implies that

$$|\operatorname{Im} a(u, u)| \leq C_3 C_4^{-1} \operatorname{Re} a(u, u), \quad u \in H. \quad (2.3.5)$$

*Proof.* With inspiration in (2.3.5), let  $\theta_0 \in ]0, \pi/2]$  be given by  $\cot \theta_0 = C_3 C_4^{-1}$ . We shall show that the sector  $\Sigma := \{0\} \cup \{\lambda \in \mathbb{C} \mid |\arg \lambda| < \frac{\pi}{2} + \theta\}$ , for each choice of  $\theta \in ]0, \theta_0[$ , is contained in the resolvent set for  $-A$ . For all  $\lambda \in \Sigma$ , we show that  $(\lambda I + A)$  is bijective, and that  $(\lambda I + A)^{-1} \in \mathbb{B}(V^*)$  with an appropriate bound on the operator norm. Note that, for each  $\lambda \in \Sigma$ , we obtain that  $\operatorname{Re} e^{i\delta} \lambda \geq 0$ , by taking  $\delta = \pm\theta$  or  $\delta = 0$ .

Therefore, for  $u \in D(A)$ ,

$$\operatorname{Re}(e^{i\delta}(A + \lambda)u \mid u) \geq \operatorname{Re}(e^{i\delta}a(u, u)) = \cos \delta \operatorname{Re} a(u, u) - \sin \delta \operatorname{Im} a(u, u). \quad (2.3.6)$$

Since,

$$0 \leq \sin \theta = \begin{cases} \sin \delta, & \text{for } \delta = \theta \\ -\sin \delta, & \text{for } \delta = -\theta, \end{cases} \quad (2.3.7)$$

multiplying (2.3.5) by  $\sin \theta$  yields, for  $\delta \in \{\pm\theta\}$ , that

$$-\sin \delta \operatorname{Im} a(u, u) \geq -C_3 C_4^{-1} \sin \theta \operatorname{Re} a(u, u), \quad (2.3.8)$$

and (2.3.6) becomes

$$\operatorname{Re}(e^{i\delta}(A + \lambda)u \mid u) \geq (\cos \theta - C_3 C_4^{-1} \sin \theta) \operatorname{Re} a(u, u). \quad (2.3.9)$$

Now,  $C_\theta := C_4 \cos \theta - C_3 \sin \theta > 0$  if and only if  $\cot \theta > C_3 C_4^{-1} = \cot \theta_0$ , and since  $\cot \theta$  is decreasing on  $]0, \pi[$ ,  $V$ -ellipticity yields

$$\operatorname{Re}(e^{i\delta}(A + \lambda)u \mid u) \geq C_\theta \|u\|_V^2, \quad \text{for } \delta \in \{0, \pm\theta\}, \quad (2.3.10)$$

obviously holding for  $\delta = 0$  too. Whence,  $m(e^{i\delta}(A + \lambda)) \geq C_\theta C_1 C_2^{-1} > 0$ . Similarly, by using that  $A^*$  is defined from the form  $a^*(v, w)$  (cf. Section 2.1), we get that  $m((e^{i\delta}(A + \lambda))^*) = m(e^{-i\delta}(A^* + \bar{\lambda})) \geq C_\theta C_1 C_2^{-1} > 0$ . So  $0 \in \rho(e^{i\delta}(A + \lambda))$  by [Gru09, Lemma 12.9]. In particular  $e^{i\delta}(A + \lambda)$ , and therefore  $A + \lambda$ , is bijective.

Furthermore, (2.3.10) implies

$$|\lambda|(u \mid u) \leq |((A + \lambda)u \mid u)| + |a(u, u)| \leq |((A + \lambda)u \mid u)| + C_3 \|u\|_V^2 \quad (2.3.11)$$

$$\leq (1 + C_3 C_\theta^{-1}) |((A + \lambda)u \mid u)|. \quad (2.3.12)$$

---

This yields that, for all  $u \in D(A)$ ,  $\lambda \in \Sigma$ ,

$$\|u\| \leq |\lambda|^{-1}(1 + C_3 C_\theta^{-1})\|(A + \lambda)u\|, \quad (2.3.13)$$

which implies (2.3.3) for  $-A$ . With minor modifications, a similar bound can be shown for  $-A^*$ .  $\square$

**Remark 2.3.4.** For  $u \in D(A)$ , inequality (2.3.5) implies that both  $\sigma(A)$  and  $\nu(A)$  are contained in the sector  $\{\lambda \in \mathbb{C} \mid |\arg \lambda| < \pi/2 - \theta_0\}$ . Whence,

$$\sigma(-A), \nu(-A) \subseteq \mathbb{C} \setminus \Sigma. \quad (2.3.14)$$

◦

We now show that  $-\mathcal{A}$  (as well as  $-\mathcal{A}'$ ) generates an analytic semigroup on  $V^*$ , which is an extension of the semigroup generated by  $A$  on  $H$ . The proof is similar to that of Lemma 2.3.3, but some changes are needed since we do not consider a fixed inner product on  $V^*$ , rather, we consider the scalar product on  $V$  and  $V^*$ .

**Lemma 2.3.5.** Let  $\mathcal{A}$  be densely defined in  $V^*$  by  $\langle \mathcal{A}v, \tilde{v} \rangle = a(v, \tilde{v})$  for all  $\tilde{v} \in V$ , with  $a(\cdot, \cdot)$  a bounded  $V$ -elliptic sesquilinear form on  $V$  as in (2.1.13) and let  $\mathcal{A}'$  be as defined there. Then  $-\mathcal{A}$  and  $-\mathcal{A}'$  generate analytic semigroups on  $V^*$ .

*Proof.* We choose  $\theta$  and  $\delta$  as in the proof of Lemma 2.3.3, and consider the sector  $\Sigma := \{\lambda \in \mathbb{C} \mid |\arg \lambda| < \frac{\pi}{2} + \theta\} \cup \{0\}$ . As before  $\operatorname{Re} e^{i\delta} \lambda \geq 0$  for  $\lambda \in \Sigma$ , and we have the inequality, for  $u \in V$ ,

$$\operatorname{Re} \langle e^{i\delta}(\mathcal{A} + \lambda)u, u \rangle = \operatorname{Re}(e^{i\delta}a(u, u)) + \operatorname{Re}(e^{i\delta}\lambda)\|u\|^2 \geq C_\theta\|u\|_V^2, \quad (2.3.15)$$

using that  $\langle u, v \rangle = (u \mid v)$  when  $u \in H$ . This implies that  $\mathcal{A} + \lambda$  is injective. We need to show that it is in fact bijective. By the previous lemma  $\lambda \in \rho(A)$ , so  $R(A + \lambda I) = H$ , which is a dense subset of  $V^*$ . Now, since  $\mathcal{A} + \lambda I$  is injective and extends  $A + \lambda I$ ,  $R(\mathcal{A} + \lambda I)$  contains  $R(A + \lambda I)$ , and hence is dense in  $V^*$ .

We now show surjectivity of  $\mathcal{A} + \lambda$  by showing that  $R(\mathcal{A} + \lambda)$  is closed. Let  $v_k$  be a sequence in  $R(\mathcal{A} + \lambda)$  that converges to  $v$ . Then, there exists  $u_k$  in  $V$  so that  $v_k = (\mathcal{A} + \lambda)u_k$ . By (2.3.15),  $u_k$  is a Cauchy sequence in  $V$ , and hence converges, but then  $u_k$  converges in  $V^*$ . Denoting the limit by  $u$ , we have that  $u_k \rightarrow u$  in  $V^*$  and  $(\mathcal{A} + \lambda)u_k \rightarrow v$  in  $V^*$ . Since  $\mathcal{A} + \lambda$  is closed,  $v = (\mathcal{A} + \lambda)u$ , and hence  $v \in R(\mathcal{A} + \lambda)$ .

We now show the bound on the resolvent. Note that, by (2.3.15), we have the following generalization of (2.3.11) for all  $u, w \in V$ , with  $u \neq 0$ ,

$$\begin{aligned} |\lambda| |\langle u, w \rangle| &\leq |\langle (\mathcal{A} + \lambda)u, w \rangle| + |a(u, w)| \leq |\langle (\mathcal{A} + \lambda)u, w \rangle| + C_3 \|u\|_V \|w\|_V \\ &\leq |\langle (\mathcal{A} + \lambda)u, w \rangle| + C_3 C_\theta^{-1} \|w\|_V |\langle (\mathcal{A} + \lambda)u, u \rangle| \|u\|_V^{-1}. \end{aligned} \quad (2.3.16)$$

Let  $\|w\|_V = 1$ , then

$$|\lambda| |\langle u, w \rangle| \leq (1 + C_3 C_\theta^{-1}) \|(\mathcal{A} + \lambda)u\|_{V^*}. \quad (2.3.17)$$

Taking the supremum over such  $w$  yields

$$\|u\|_{V^*} \leq |\lambda|^{-1} (1 + C_3 C_\theta^{-1}) \|(\mathcal{A} + \lambda)u\|_{V^*}. \quad (2.3.18)$$

Therefore,  $\lambda \in \rho(-\mathcal{A})$  with the bound needed.

With only minor modifications to the proof, it can be shown that  $-\mathcal{A}'$  generates an analytic semigroup on  $V^*$ .  $\square$

To avoid confusion with the notation  $T$  for the terminal time, we denote the semigroup on  $H$  generated by  $-A$  by  $e^{-tA}$ , even though it is not to be understood as an exponential function having a power series expansion, as  $-A$  is not necessarily bounded. The notation, however, is consistent with standard functional calculus, cf. (2.3.21) below. Likewise, we denote by  $e^{-t\mathcal{A}}$  the semigroup generated by  $-\mathcal{A}$  on  $V^*$ .

Since  $e^{-t\mathcal{A}}$  is a semigroup,  $e^{-t\mathcal{A}} \in \mathbb{B}(V^*)$  and it has an adjoint operator in  $\mathbb{B}(V)$ , which we denote by  $(e^{-t\mathcal{A}})^*$ . Since  $V^*$  is a reflexive space, it is a result ([Paz83, Corollary 1.10.6]) that the adjoint semigroup is generated by the adjoint of the generator. Denoting the adjoint of the generator  $-\mathcal{A}$  by  $-\mathcal{B}$ , i.e.

$$\langle -\mathcal{A}w, v \rangle = \langle w, -\mathcal{B}v \rangle, \quad \text{for all } v \in D(\mathcal{B}), w \in V, \quad (2.3.19)$$

we have that  $(e^{-t\mathcal{A}})^* = e^{-t\mathcal{B}}$  in  $\mathbb{B}(V)$ . Similarly,  $(e^{-tA})^* = e^{-tA^*}$  in  $\mathbb{B}(H)$ . That  $e^{-tA^*}$  is an analytic semigroup follows from Lemma 2.3.3, and it is an exercise to show that also  $e^{-t\mathcal{B}}$  is an analytic semigroups since  $e^{-tA}$  are so.

We now exploit that the sector  $\Sigma$  can be chosen equal for  $A$  and  $\mathcal{A}$ , which results from the proofs above, to show that  $e^{-t\mathcal{A}}$  extends  $e^{-tA}$ .

**Proposition 2.3.6.** *Let  $A$  be a Lax-Milgram operator induced by the triple  $(V, H, a(\cdot, \cdot))$ , where  $a(\cdot, \cdot)$  is a bounded,  $V$ -elliptic sesquilinear form, and  $V \subseteq H$  algebraically, topologically and densely, and let  $\mathcal{A}$  be the extension of  $A$  to a mapping  $V \rightarrow V^*$ . Then, the semigroups generated by  $-A$  and  $-\mathcal{A}$  agree on  $H$ .*

*Proof.* From the identities  $(\mathcal{A} + \lambda I)(\mathcal{A} + \lambda I)^{-1} = I_{V^*}$  and  $(\mathcal{A} + \lambda I)^{-1}(\mathcal{A} + \lambda I) = I_V$ , it follows, since  $A = \mathcal{A}|_{\mathcal{A}^{-1}(H)}$ , cf. Section 2.1, that

$$(A + \lambda I)^{-1} = (\mathcal{A} + \lambda I)^{-1}|_H. \quad (2.3.20)$$

Now, from [Paz83, Theorem 1.7.7], we have that

$$e^{-tA} = \frac{1}{2\pi i} \int_{\Gamma} e^{t\mu} (-A - \mu)^{-1} d\mu, \quad (2.3.21)$$

where  $\Gamma$  is a smooth curve in the sector  $\Sigma$  from Lemma 2.3.3, running from  $\infty e^{-i\kappa}$  to  $\infty e^{i\kappa}$  with  $\pi/2 < \kappa < \pi/2 + \theta$ , and the integral converges to  $e^{-tA}$  in the uniform operator topology. Note that the integral is well-defined since  $\mu \in \rho(A)$ . Since the integral as a consequence converges strongly, i.e. for all  $v \in V^*$

$$e^{-tA}v = \frac{1}{2\pi i} \int_{\Gamma} e^{t\mu} (-A - \mu)^{-1} v d\mu, \quad (2.3.22)$$

it follows by (2.3.20) that  $e^{-tA}|_H = e^{-tA}$ , i.e.  $e^{-tA}$  extends  $e^{-tA}$ .  $\square$

**Remark 2.3.7.** *Note that the range of the semigroup decreases as  $t$  increases, since, by the semigroup property,*

$$R(e^{-(t+\varepsilon)A}) = R(e^{-tA}e^{-\varepsilon A}) \subseteq R(e^{-tA}). \quad (2.3.23)$$

$\circ$

Next, we make some additional remarks on the action of a semigroup on a vector function. We start by proving the following technical lemma regarding the measurability of the semigroup acting on a vector function (undoubtedly folklore).

**Lemma 2.3.8.** *Let  $A$  be a generator of an analytic semigroup on Banach space  $W$  with anti-dual  $W^*$ . For  $f \in L_1(0, T; W)$ , and  $t \in [0, T]$ , we have that  $e^{(t-\cdot)A}f \in L_1(0, t, W)$ , and that  $\langle \eta, e^{(t-\cdot)A}f(\cdot) \rangle_{W^*, W} \in L_1(0, t)$ , for  $\eta \in W^*$ .*

*Proof.* We start by showing that  $e^{(t-\cdot)A}f$  is strongly measurable, i.e. that there is a sequence of simple functions converging pointwise to  $e^{(t-s)A}f(s)$  for a.e.  $s \in [0, t]$  (see e.g. [RS80, page 116]).

Since  $f \in L_2(0, t; W)$ ,  $f$  is strongly measurable. Let  $\{f_n\}$  be a sequence of simple functions, with  $f_n$  taking the values  $y_1^{(n)}, \dots, y_{m_n}^{(n)}$ , such that  $\|f_n(s) - f(s)\|_W \rightarrow 0$  for a.e.  $s < t$ . For such an  $s$ ,  $e^{(t-s)A} \in \mathbb{B}(W)$  implies that, for  $\varepsilon > 0$ , there is a  $N$  such that for  $n \geq N$ ,

$$\|e^{(t-s)A}f_n(s) - e^{(t-s)A}f(s)\|_W < \frac{\varepsilon}{2}. \quad (2.3.24)$$

For  $s < t$ ,  $e^{(t-s)A}$  is continuous in  $\mathbb{B}(W)$ . Hence it is uniformly continuous in  $\mathbb{B}(W)$  on the compact subinterval  $[0, t - 2^{-n}]$ . So, for  $\varepsilon > 0$ , there exist  $K(n)$  so that, for  $|s_1 - s_2| < t \cdot 2^{-k}$ ,  $k \geq K(n)$ ,

$$\|e^{(t-s_1)A} - e^{(t-s_2)A}\|_{\mathbb{B}(W)} \leq \frac{\varepsilon}{2(1 + \|y_1\|_W + \dots + \|y_{m_n}^{(n)}\|_W)}. \quad (2.3.25)$$

Here we may choose  $K(n)$  to be increasing in  $n$ .

We split  $[0, t]$  into  $2^{K(n)}$  subintervals of length  $t \cdot 2^{-K(n)}$ , and note that each  $s < t$  lies precisely in one of these, denoted by  $[l_s \cdot t \cdot 2^{-K(n)}, (l_s + 1) \cdot t \cdot 2^{-K(n)}]$ . On this subinterval, we approximate the semigroup with the value at  $l_s \cdot t \cdot 2^{-K(n)}$ . Hence, let

$$g_n(s) := \begin{cases} e^{(t-l_s \cdot t \cdot 2^{-K(n)})A} f_n(s), & \text{on } [0, t - 2^{-n}] \\ 0, & \text{for } s > t - 2^{-n}. \end{cases} \quad (2.3.26)$$

This is a simple function, since  $f_n$  is so, and there is a finite number of subintervals.

To show that  $g_n(s)$  converges to  $e^{(t-s)A} f(s)$  a.e, let  $\varepsilon > 0$  be given, and consider an  $s < t$  for which we have pointwise convergence, i.e.  $\|f_n(s) - f(s)\|_W \rightarrow 0$ . We may choose  $N_1$  so large that  $s \in [0, t - 2^{-N_1}]$ , and  $N_2$  so large that (2.3.24) holds. Let  $N = \max\{N_1, N_2\}$ . Since  $K(n) \geq K(N)$  for  $n \geq N$ , we get by (2.3.24), and (2.3.25), that

$$\begin{aligned} \|g_n(s) - e^{(t-s)A} f(s)\|_W &\leq \|g_n(s) - e^{(t-s)A} f_n(s)\|_W \\ &\quad + \|e^{(t-s)A} f_n(s) - e^{(t-s)A} f(s)\|_W \\ &\leq \varepsilon(2(1 + \|\tilde{y}_1\|_W + \dots + \|\tilde{y}_{m_n}^{(n)}\|_W))^{-1} \|f_n(s)\|_W + \frac{\varepsilon}{2} < \varepsilon. \end{aligned} \quad (2.3.27)$$

This shows the pointwise convergence a.e. on  $[0, t]$ . Hence  $e^{-(t-\cdot)A} f$  is strongly measurable.

Now  $e^{(t-\cdot)A} f \in L_1(0, T; W)$ , for by the boundedness of  $e^{(t-\cdot)A}$ , we see that there exist constants  $M \geq 1$ ,  $\omega \geq 0$  (cf. (2.3.1)) such that

$$\|e^{(t-\cdot)A} f\|_{L_1(0, T; W)} \leq M e^{\omega t} \|f\|_{L_1(0, T; W)}. \quad (2.3.28)$$

Since strong measurability implies weak measurability (see e.g. [RS80, page 116]), we see that  $\langle \eta, e^{(t-\cdot)A} f \rangle_{W^*, W} \in L_1(0, t)$  by majorizing with  $\|e^{(t-s)A} f(s)\|_W \|\eta\|_{W^*}$ .  $\square$

We now show a Leibniz rule for differentiation of a semigroup acting on a vector function.

**Proposition 2.3.9.** *Let  $A$  be the generator of an analytic semigroup on a Banach space  $W$ . Then, for  $w \in H^1(0, T; W)$ ,  $T > 0$ , we have in  $\mathcal{D}'(0, T; W)$ , that*

$$\partial_s(e^{(T-s)A} w(s)) = (-A)e^{(T-s)A} w(s) + e^{(T-s)A} \partial_s w(s). \quad (2.3.29)$$

*Proof.* First, we show (2.3.29) for  $w \in C^1(0, T; W)$ , and later extend it to arbitrary  $w \in H^1(0, T; W)$  by a density argument. Let  $w \in C^1(0, T; W)$  so that we have  $w(s+h) - w(s) = h\partial_s w(s) + o(h)$  in the limit  $h \rightarrow 0$ . For  $s < T$ , the semigroup is differentiable (cf. Theorem 2.3.2), so  $e^{(T-(s+h))A}x - e^{(T-s)A}x = h(-A)e^{(T-s)A}x + o(h)$ ,  $x \in W$ . Since  $e^{(T-s)A}$  is analytic for  $0 < s < t$ ,  $(-A)e^{(T-s)A}o(1) = o(1)$  for  $h \rightarrow 0$ , by the boundedness of  $(-A)e^{(T-s)A}$  in (2.3.4). Following the lines of the proof of Lemma 2.1.2, we get, by adding and subtracting appropriate terms,

$$\begin{aligned} e^{(T-(s+h))A}w(s+h) - e^{(T-s)A}w(s) &= (h(-A)e^{(T-s)A} + o(h))(h\partial_s w(s) + o(h)) \\ &\quad + (h(-A)e^{(T-s)A} + o(h))w(s) \\ &\quad + e^{(T-s)A}(h\partial_s w(s) + o(h)) \\ &= h((-A)e^{(T-s)A}w(s) + e^{(T-s)A}\partial_s w(s)) \\ &\quad + o(h), \end{aligned} \tag{2.3.30}$$

which shows (2.3.29) for  $w \in C^1(0, T; W)$ .

As in the proof of Lemma 2.1.2, for  $w \in H^1(0, T; W)$ , there is a sequence  $\{w_k\} \in C^1(0, T; W)$  such that  $w_k \rightarrow w$  in  $L_2(0, T; W)$ , and  $w'_k \rightarrow w'$  in  $L_{2,\text{loc}}(0, T; W)$ .

For  $\phi \in D(0, T)$ , by Lemma 2.3.8, the Bochner inequality (see e.g. [Eva08]), and the boundedness of the semigroup in (2.3.1), we have that

$$\begin{aligned} \left\| \int_0^T e^{(T-s)A}(w(s) - w_k(s))\phi(s) \, ds \right\|_W &\leq \int_{\text{supp } \phi} \|e^{(T-s)A}(w(s) - w_k(s))\phi(s)\|_W \, ds \\ &\leq C\|w(s) - w_k(s)\|_{L_2(0, T; W)}, \end{aligned} \tag{2.3.31}$$

with  $C = M(\int_{\text{supp } \phi} |\phi(s)e^{(T-s)A}|^2 \, ds)^{1/2}$ , where  $M$  is the constant in (2.3.1).

Hence,  $e^{(T-s)A}w_k \rightarrow e^{(T-s)A}w$  in  $\mathcal{D}'(0, T; W)$ , and  $\partial_s(e^{(T-s)A}w) = \lim_{k \rightarrow \infty} (\partial_s(e^{(T-s)A}w_k))$ .

Moreover, for  $\varepsilon > 0$  chosen such that  $\text{supp } \phi \subseteq ]\varepsilon, T - \varepsilon[$ ,

$$\left\| \int_{\text{supp } \phi} e^{(T-s)A}(w'(s) - w'_k(s))\phi(s) \, ds \right\|_W \leq c\|w'(s) - w'_k(s)\|_{L_2(\varepsilon, T-\varepsilon; W)}. \tag{2.3.32}$$

Furthermore, since  $(-A)e^{(T-s)A}$  is bounded for  $s < T$ , cf. (2.3.4),

$$\left\| \int_0^T (-A)e^{(T-s)A}(w(s) - w_k(s))\phi(s) \, ds \right\|_W \leq \tilde{C}\|w(s) - w_k(s)\|_{L_2(0, T; W)}, \tag{2.3.33}$$

with  $\tilde{C} = (\int_{\text{supp } \phi} |\frac{c\phi(s)}{(T-s)}|^2 \, ds)^{1/2}$ . Note that  $s \neq T$  for  $s \in \text{supp } \phi$ . Since formula (2.3.29) obviously holds for  $w_k$ , for all  $k$ , it follows from (2.3.31) - (2.3.33), that

---

formula (2.3.29) extends to the arbitrary  $w \in H^1(0, T; W)$  by passing to the limit.  $\square$

**Corollary 2.3.10.** *If furthermore  $w(s) \in D(A)$  for all  $s \in [0, T]$ , then*

$$\partial_s(e^{(T-s)A}w(s)) = e^{(T-s)A}(-A)w(s) + e^{(T-s)A}\partial_s w(s). \quad (2.3.34)$$

*Proof.* In particular,  $w \in H^1(0, T; W)$ , so (2.3.29) holds. Now,  $w(s) \in D(A)$  for all  $s$  implies, by [Paz83, Theorem 1.2.4], that  $(-A)e^{(T-s)A}w(s) = e^{(T-s)A}(-A)w(s)$ , which completes the proof.  $\square$

## 2.4. A representation of the solution

In this section we combine the theories introduced in the two previous sections, and have the following theorem.

**Theorem 2.4.1.** *Let  $V \subseteq H \subseteq V^*$  algebraically, topologically and densely. Let  $A$  denote a Lax-Milgram operator defined from the triple  $(H, V, a(\cdot, \cdot))$ , with  $a(\cdot, \cdot)$  a bounded, and  $V$ -elliptic sesquilinear form on  $V$ , and  $\mathcal{A} : V \rightarrow V^*$  the extension of  $A$ . Consider the initial value problem*

$$\begin{cases} u' + \mathcal{A}u = f & \text{in } \mathcal{D}'(0, T; V^*), \\ u(0) = u_0 & \text{in } H. \end{cases} \quad (2.4.1)$$

*For  $u_0 \in H$ , and  $f \in L_2(0, T; V^*)$ , we have the following representation of the solution,  $u \in X$ , of (2.4.1)*

$$u(t) = e^{-tA}u_0 + \int_0^t e^{-(t-s)A}f(s) ds. \quad (2.4.2)$$

First note, that Theorem 2.2.1 gives the existence and uniqueness of a solution. Hence, we need only prove that the solution is on the form (2.4.2).

The strategy of proof is to use semigroup theory to define a function which, as an element of  $V^*$  acting on  $V$ , has a suitable distributional derivative, in that we want to be able to use Lemma 2.2.2. This is the contents of the following lemma.

**Lemma 2.4.2.** *When  $u(t)$  denotes the solution of (2.4.1), let*

$$g(s) := e^{-(t-s)A}u(s), \quad 0 \leq s \leq t. \quad (2.4.3)$$

*Then, in the sense of scalar distributions,*

$$\partial_s \langle g(s), \eta \rangle = \langle e^{-(t-s)A}f(s), \eta \rangle, \quad \eta \in V. \quad (2.4.4)$$

---

*Proof.* We see that  $g \in L_1(0, T; V^*)$  by Lemma 2.3.8, as  $u \in L_2(0, T; V^*) \subseteq L_1(0, T; V^*)$ . Furthermore,  $\langle g(s), \eta \rangle \in L_1(0, T)$  for  $\eta \in V$ , so it makes sense to derive it in the scalar distributional sense. Also,  $f \in L_2(0, T; V^*) \subseteq L_1(0, T; V^*)$  by assumption, so  $e^{-(t-s)\mathcal{A}}f(s) \in L_1(0, T; V^*)$ , and  $\langle e^{-(t-s)\mathcal{A}}f(s), \eta \rangle$  is in  $L_1(0, T)$  by Lemma 2.3.8. Therefore, both sides of (2.4.4) make sense.

To avoid introducing a singularity at  $s = t$  by direct application of (2.3.4), note that, since  $e^{-t\mathcal{B}}$  is analytic, one can use Lemma 2.1.2 for  $s \in ]0, t - \varepsilon[$ ,  $\varepsilon > 0$ , to get,

$$\partial_s \langle u(s), e^{-(t-s)\mathcal{B}}\eta \rangle = \langle u'(s), e^{-(t-s)\mathcal{B}}\eta \rangle + \langle u(s), \mathcal{B}e^{-(t-s)\mathcal{B}}\eta \rangle. \quad (2.4.5)$$

Since  $u$  solves (2.2.1), we conclude that

$$\partial_s \langle g(s), \eta \rangle = \langle e^{-(t-s)\mathcal{A}}u'(s), \eta \rangle + \langle e^{-(t-s)\mathcal{A}}\mathcal{A}u(s), \eta \rangle = \langle e^{-(t-s)\mathcal{A}}f(s), \eta \rangle. \quad (2.4.6)$$

Since  $\varepsilon > 0$  is arbitrary, this shows (2.4.4) in  $\mathcal{D}'(0, T)$ . □

We now prove the representation formula.

*Proof of Theorem 2.4.1.* Let  $u$  denote the solution of (2.4.1), the existence and uniqueness of which follows from Theorem 2.2.1.

To apply Lemma 2.2.2, let  $g(s)$  in (2.4.3) play the role of  $u$  in that lemma with  $B = V^*$ . Then it follows from the implication (2.2.5)  $\Rightarrow$  (2.2.3), that (2.4.4) implies

$$g(t) = g(0) + \int_0^t e^{-(t-s)\mathcal{A}}f(s) \, ds, \quad \text{a.e.} \quad (2.4.7)$$

By definition of  $g$ , and using Proposition 2.3.6 to conclude that  $e^{-t\mathcal{A}}u_0 = e^{-t\mathcal{A}}u_0$ , we have

$$u(t) = e^{-t\mathcal{A}}u_0 + \int_0^t e^{-(t-s)\mathcal{A}}f(s) \, ds \quad \text{a.e.} \quad (2.4.8)$$

This is the needed representation of  $u$  □

## 2.5. Notes

Let us elaborate on the gain from joining the two methods. In semigroup theory, in general, it is easy to see that the mild solution in (2.4.2) is the only possible



---

solution to (2.4.1). However, in order to prove existence of a solution, one needs to assume Hölder continuity of  $f : [0, T] \rightarrow H$  of some appropriate order (see e.g. [Paz83, Chapter 4 & 7]). In combination with functional analysis (cf. [Tem01] and the minor generalizations in Theorem 2.2.1 above), especially the use of vector distributions, we have proven that the mild solution is indeed a solution, in the sense of distributions, under much weaker conditions on  $f$ , namely we need only assume that  $f \in L_2(0, T; V^*)$ . It is however crucial that the generator  $A$  is a Lax-Milgram operator.

### 3. Well-posedness of the final value evolution problem

In this chapter, we show that the following final value evolution problem

$$\begin{cases} u'(t) + \mathcal{A}u(t) = f \\ u(T) = u_T \end{cases} \quad (3.0.1)$$

is *well-posed* when the final data lies in an appropriate space, identified below. This is essentially done by comparison with the corresponding initial value problem treated in the previous chapter, cf (2.0.1). Specifically, we show, in Section 3.1, that there exists a bijective correspondence between the initial states of (2.0.1) and the terminal states of (3.0.1). This bijective correspondence is used to prove a well-posedness result in Section 3.2.

#### 3.1. A bijection between initial and terminal data

Let  $A$  denote a  $V$ -elliptic Lax-Milgram operator, and  $\mathcal{A} : V \rightarrow V^*$  its extension, cf. Section 2.1. In this section, we study  $u$  at terminal time  $T$ , (cf. (2.4.2)),

$$u(T) = e^{-TA}u_0 + \int_0^T e^{-(T-s)\mathcal{A}}f(s) \, ds. \quad (3.1.1)$$

This representation is essential in what follows, since it gives a bijective correspondence between initial and terminal data, explored below.

First, introduce the following notation, which will be useful throughout the text,

$$y_f := \int_0^T e^{-(T-s)\mathcal{A}}f(s) \, ds. \quad (3.1.2)$$

Note that by Lemma 2.3.8,  $f \mapsto y_f$  is a well-defined linear map. By (2.3.1), we have  $\|e^{-(T-s)\mathcal{A}}\|_{\mathbb{B}(V^*)} \leq M e^{-\omega(T-s)}$ , and it follows by the Cauchy-Schwartz inequality that

$$\|y_f\|_{V^*} \leq \|M e^{-\omega(T-s)}\|_{L_2(0,T)} \|f\|_{L_2(0,T;V^*)} \leq \frac{M}{\sqrt{\omega}} \|f\|_{L_2(0,T;V^*)}. \quad (3.1.3)$$

However, the value of  $y_f$  is actually in the smaller space  $H$ .

---

**Lemma 3.1.1.** *When  $f \in L_2(0, T; V^*)$ , then  $\|y_f\|_H \leq c\|f\|_{L_2(0, T; V^*)}$ .*

*Proof.* Since  $y_f = u(T)$  with  $u$  the solution of (2.4.1) for  $u_0 = 0$ , Corollary 2.2.6 gives

$$\|y_f\|_H \leq \sup_{t \in [0, T]} \|u(t, x)\|_H \leq c\|f\|_{L_2(0, T; V^*)}, \quad (3.1.4)$$

improving (3.1.3). □

Often, the range of  $f \mapsto y_f$  is quite large, as one can show by an explicit construction. This is the contents of the following lemma.

**Proposition 3.1.2.** *Assume that  $H$  has an orthonormal basis of eigenvectors for  $A$ , then the mapping  $y_f : L_2(0, T; V^*) \rightarrow H$  fulfills that  $V \subseteq R(y_f)$ .*

*Proof.* Let  $v \in V$  with  $v = \sum_{j=1}^{\infty} \alpha_j e_j = \sum_{j=1}^{\infty} (v | e_j) e_j$ . We shall choose  $f \in L_2(0, T; V^*)$  with  $f(t) = \sum_{j=1}^{\infty} \beta_j e_j = \sum_{j=1}^{\infty} \langle f(t), e_j \rangle e_j$  such that  $y_f = v$ . Let  $\beta_j(t) = \mathbb{1}_{[0, T]} \beta_j$ . Since  $e^{-(T-s)A} e_j = e^{-(T-s)\lambda_j} e_j$ , cf. Remark 3.1.4 below, we have

$$y_f = \int_0^T e^{-(T-s)A} f(s) ds = \sum_{j=1}^{\infty} \beta_j \int_0^T e^{-(T-s)\lambda_j} e_j ds, \quad (3.1.5)$$

which, when  $y_f = v$ , implies that

$$\beta_j = \alpha_j \lambda_j (1 - e^{-T\lambda_j})^{-1}. \quad (3.1.6)$$

Now, by Fact 1, Fact 2, and since we may order the eigenvalues increasingly such that  $\lambda_j \geq \lambda_1$  for all  $j$ ,

$$\|f\|_{L_2(0, T; V^*)}^2 = \int_0^T \|f(t)\|_{V^*}^2 dt = \int_0^T \sum_{j=1}^{\infty} \lambda_j^{-1} \beta_j^2 dt = T \sum_{j=1}^{\infty} \lambda_j^{-1} \alpha_j^2 \lambda_j^2 (1 - e^{-T\lambda_j})^{-2} \quad (3.1.7)$$

$$\leq T(1 - e^{-T\lambda_1})^{-2} \sum_{j=1}^{\infty} \lambda_j \alpha_j^2 < \infty. \quad (3.1.8)$$

Hence, we may choose  $f$  so that  $y_f = v$ . □

It is mathematical folklore that the semigroup is injective at each  $t \geq 0$ . Even when the sesquilinear form is  $t$ -dependent, we have that  $u_0$  is uniquely determined from  $u_T$  (the interested reader is referred to [LM60]). In lack of a reference when the sesquilinear form is not  $t$ -dependent, we include a straightforward proof below.

---

**Proposition 3.1.3.** *Let  $e^{-zA}$  denote a semigroup on  $H$  which is analytic in the sector  $S := \{z \in \mathbb{C} \mid |\arg z| < \delta\}$  for  $\delta \in ]0, \pi[$ . Then  $e^{-zA}$  is injective for  $z \in S$ .*

*Proof.* To show that the null space  $Z(e^{-zA})$  is trivial, we let  $u_0 \in H$  and a  $z_0 \in S$  be such that

$$e^{-z_0 A} u_0 = 0. \quad (3.1.9)$$

Set  $f : z \mapsto e^{-zA} u_0$ , and  $g_v : z \mapsto (f(z) \mid v)$ , for arbitrary  $v \in H$ . Then, analyticity of  $e^{-zA}$  and the chain rule imply that  $f$  and  $g_v$  are analytic with  $f'(z) = (e^{-zA})' u_0$ , and  $g'_v(z) = (f'(z) \mid v)$ ,  $z \neq 0$  (analogously to Lemma 2.1.1). Hence,  $g_v$  has, in an some open ball  $B(z_0, r) \subseteq S$ , the unique Taylor expansion

$$g_v(z) = \sum_{n=0}^{\infty} \frac{1}{n!} g^{(n)}(z_0) (z - z_0)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \left( f^{(n)}(z_0) \mid v \right) (z - z_0)^n. \quad (3.1.10)$$

By (3.1.9), and the standard formula for derivatives of analytic semigroups (See e.g. [Paz83, Lemma 2.4.2]), which generalizes in a straightforward way to yield

$$f^{(n)}(z_0) = (-A)^n e^{-z_0 A} u_0 = 0 \quad \text{for all } n \geq 0, \quad (3.1.11)$$

so that  $g_v \equiv 0$  on  $B(z_0, r)$ , and hence on  $S$  by the unique analytic extension.

This implies that  $f \equiv 0$  on  $S$ , since  $f(z_1) \neq 0$  would contradict  $g_v \equiv 0$  for  $v = f(z_1)$ . So, since  $e^{-tA}$  is a strongly continuous semigroup,

$$u_0 = \lim_{t \rightarrow 0} e^{-tA} u_0 = \lim_{t \rightarrow 0} f(t) = 0. \quad (3.1.12)$$

□

**Remark 3.1.4.** *In the particular case where  $H$  has an orthonormal basis of eigenvectors for  $A$ , the injectivity follows explicitly by computations, which we include in the following. Let  $\{e_j\}$  be an orthonormal basis of eigenvectors of  $A$  such that  $Ae_j = \lambda_j e_j$ . It is immediate that  $e^{-t\lambda_j} e_j$  and  $e^{-tA} e_j$  both solves*

$$\begin{cases} x' + Ax = 0 \\ x(0) = e_j, \end{cases} \quad (3.1.13)$$

*so that  $e^{-tA} e_j = e^{-t\lambda_j} e_j$ , for all  $t \in [0, T]$ , by uniqueness of the solution. Whence, if  $e^{-tA} v = 0$  for some  $v \in H$ ,  $t \in [0, T]$ , the boundedness of  $e^{-tA}$  gives*

$$0 = e^{-tA} v = \sum_j (v \mid e_j) e^{-tA} e_j = \sum_j (v \mid e_j) e^{-t\lambda_j} e_j, \quad (3.1.14)$$

*implying that  $\sum_j |(v \mid e_j)|^2 e^{-2t\lambda_j} = 0$ , and  $v \equiv 0$ . Hence,  $e^{-tA}$  is invertible. ◻*

We use the symbol  $e^{tA}$  to denote the inverse of  $e^{-tA}$ . This is an abuse of the semigroup notation because  $e^{tA}$  is not in general a group, but it is convenient for our purposes (although it requires some diligence). To emphasize this unconventional choice, it is recalled once more that, for the Lax-Milgram operator  $A$ , the analytic semigroup generated by  $-A$  is written  $e^{-tA}$ , and that

$$e^{tA} = (e^{-tA})^{-1}. \quad (3.1.15)$$

It maps  $R(e^{-tA})$  onto  $D(e^{-tA}) = H$ , and has the following properties

**Lemma 3.1.5.** *The inverse in (3.1.15) is a bounded operator from its domain  $D(e^{tA}) = R(e^{-tA})$  into  $H$  with  $\|e^{tA}\|_{\mathbb{B}(D(e^{tA}), H)} < 1$ , and  $D(e^{tA})$  decreases in  $t$ , i.e., for  $t' \geq t$ ,*

$$D(e^{t'A}) \subseteq D(e^{tA}). \quad (3.1.16)$$

*In analogy to the semigroup property, we have*

$$e^{(T-t)A} \supseteq e^{-tA} e^{TA}, \quad \text{for } 0 \leq t \leq T. \quad (3.1.17)$$

*Proof.* As  $\|e^{tA}x\|_H < \|x\|_{D(e^{tA})}$  when  $x \neq 0$ , it follows that  $e^{tA}$  is a bounded operator from its domain into  $H$  with

$$\|e^{tA}\|_{\mathbb{B}(D(e^{tA}), H)} < 1. \quad (3.1.18)$$

From Remark 2.3.7, it follows that  $D(e^{t'A}) \subseteq D(e^{tA})$ , for  $t' \geq t$ .

By the semigroup property

$$e^{-(T-t)A} e^{-tA} e^{TA} = e^{-TA} e^{TA} = I_{R(e^{-TA})}, \quad (3.1.19)$$

so by Remark 2.3.7, one finds

$$e^{(T-t)A} = e^{(T-t)A} I_{R(e^{-(T-t)A})} \supseteq e^{(T-t)A} I_{R(e^{-TA})} = e^{-tA} e^{TA}, \quad (3.1.20)$$

which yields (3.1.17).  $\square$

**Remark 3.1.6.** *Using the bijectiveness of the semigroup, we note that  $R(y_f) \subseteq H$  densely. Indeed, recalling from Lemma 3.1.1 that  $y_f$  in (3.1.2) is a map  $L_2(0, T; V^*) \rightarrow H$ , we see that for all  $v \in H$ , by the Bochner identity (see e.g. [Eva08]), noting that  $e^{-(T-s)\mathcal{B}}v \in D(\mathcal{B}) \subseteq V$ ,*

$$(y_f \mid v) = \int_0^T \left( e^{-(T-s)\mathcal{A}} f(s) \mid v \right) ds = \int_0^T \langle f(s), e^{-(T-s)\mathcal{B}}v \rangle_{V^*, V} ds. \quad (3.1.21)$$

*The adjoint  $y_f^* : H \rightarrow L_2(0, T; V)$  is therefore given by  $y_f^*v = e^{-(T-s)\mathcal{B}}v$  which is injective by Proposition 3.1.3, since the adjoint of an analytic semigroup, on a Hilbert space, is again an analytic semigroup. Hence,  $R(y_f) = Z(y_f^*)^\perp = \{0\}^\perp = H$ .  $\circ$*

---

**Remark 3.1.7.** *It is worth noting that  $D(e^{tA}) = R(e^{-tA}) \subseteq H$  densely for any  $0 < t < \infty$ , which is seen in much the same way as in the previous remark.  $\circ$*

The following proposition is an improvement to Remark 2.3.7 and Lemma 3.1.5 when we impose additional conditions on  $H$ .

**Proposition 3.1.8.** *When  $H$  has an orthonormal bases of eigenvectors  $\{e_j\}$  of  $A$ , with the corresponding eigenvalues  $\lambda_j \rightarrow \infty$  for  $j \rightarrow \infty$ , then there are, for any  $0 < t < T < \infty$ , strict inclusions*

$$D(e^{TA}) \subset D(e^{tA}) \subset H. \quad (3.1.22)$$

*Moreover,  $D(e^{tA})$  is the completion of  $\text{Span}\{e_j \mid j \in \mathbb{N}\}$  with respect to the graph norm on  $D(e^{tA})$ .*

*Proof.* First, we obtain that the graph norm on  $D(e^{tA})$  is given by

$$\|x\|_{D(e^{tA})}^2 = \sum_{j=1}^{\infty} (1 + e^{2\lambda_j t}) |(x \mid e_j)|^2, \quad (3.1.23)$$

by noting that

$$D(e^{tA}) = S := \left\{ x \in H \mid \sum_{j=1}^{\infty} e^{2\lambda_j t} |(x \mid e_j)|^2 < \infty \right\}. \quad (3.1.24)$$

In fact, for  $x \in S$ ,  $y = \sum_{j=1}^{\infty} e^{\lambda_j t} (x \mid e_j) e_j$  is a well defined vector in  $H$ , and since  $e^{-tA} e_j = e^{-\lambda_j t} e_j$  (cf. Remark 3.1.4), boundedness of  $e^{-tA}$  yields

$$e^{-tA} y = \sum_{j=1}^{\infty} e^{\lambda_j t} (x \mid e_j) e^{-tA} e_j = \sum_{j=1}^{\infty} (x \mid e_j) e_j = x, \quad (3.1.25)$$

so that  $x \in D(e^{tA})$ . Conversely,  $x \in D(e^{tA}) = R(e^{-tA})$  implies that there is a  $y \in H$  such that

$$x = e^{-tA} y = \sum_{j=1}^{\infty} (y \mid e_j) e^{-t\lambda_j} e_j, \quad (3.1.26)$$

so that  $e^{\lambda_j t} (x \mid e_j) = (y \mid e_j) \in \ell_2$ , i.e.  $x \in S$ .

Now, let  $t_1 < t_2$ . For  $x \in D(e^{t_2 A})$ , we immediately have, since  $\lambda_j > 0$  by  $V$ -ellipticity, that  $x \in D(e^{t_1 A})$ , since

$$\sum_{j=1}^{\infty} e^{2\lambda_j t_1} |(x \mid e_j)|^2 < \sum_{j=1}^{\infty} e^{2\lambda_j t_2} |(x \mid e_j)|^2 < \infty. \quad (3.1.27)$$

On the other hand, we may choose  $x \in D(e^{t_1 A})$  in such a way that  $x \notin D(e^{t_2 A})$ . Since  $\lambda_j \rightarrow \infty$  by assumption, we may choose a subsequence  $\{\lambda_{j_n}\}$ , such that  $\lambda_{j_n} > 2^n$  for all  $n \in \mathbb{N}$ . Now, for  $\{a_n\} \in \ell_2$ , let

$$x = \sum_{n=1}^{\infty} a_n e^{-\lambda_{j_n} t_1} e_{j_n}. \quad (3.1.28)$$

This vector is in  $D(e^{t_1 A})$  because

$$\sum_{j=1}^{\infty} e^{2\lambda_j t_1} |(x | e_j)|^2 = \sum_{n=1}^{\infty} |a_n|^2 < \infty, \quad (3.1.29)$$

whereas, e.g.  $a_n = n^{-1}$ , gives

$$\sum_{j=1}^{\infty} e^{2\lambda_j t_2} |(x | e_j)|^2 = \sum_{n=1}^{\infty} e^{2\lambda_{j_n} (t_2 - t_1)} |a_n|^2 > \sum_{n=1}^{\infty} e^{2^n (t_2 - t_1)} |a_n|^2 \geq \infty. \quad (3.1.30)$$

Hence,  $D(e^{t_2 A}) \subset D(e^{t_1 A})$  is a strict inclusion.

Furthermore,  $D(e^{tA})$  is the completion of  $\text{Span}\{e_j \mid j \in \mathbb{N}\}$  with respect to the graph norm on  $D(e^{tA})$ . Indeed, by (3.1.23), it follows that, for  $x \in D(e^{tA})$ ,

$$x - \sum_{j \leq N} (x | e_j) e_j \rightarrow 0 \text{ in } D(e^{tA}) \text{ as } N \rightarrow \infty, \quad (3.1.31)$$

which completes the proof.  $\square$

**Remark 3.1.9.** *The previous remark can be extended to the semigroup on  $V^*$  generated by  $-\mathcal{A}$ , by using Fact 2, to see that*

$$D(e^{t\mathcal{A}}) = \{w \in V^* \mid \sum_{j=1}^{\infty} e^{2\lambda_j t} |\langle w, e_j \rangle|^2 < \infty\}. \quad (3.1.32)$$

$\circ$

After this study of  $y_f$ , and  $D(e^{tA})$ , note that the mapping

$$u_0 \mapsto e^{-TA} u_0 \mapsto e^{-TA} u_0 + y_f, \quad (3.1.33)$$

is composed of the bijection  $e^{-TA}$ , and a translation by  $y_f$ , hence is bijective from  $H$  to the affine space  $R(e^{-TA}) + y_f$ , so we have the following theorem.

**Theorem 3.1.10.** *For the set of solutions  $u$  of the differential equation  $(\partial_t + \mathcal{A})u = f$  with data fixed  $f \in L_2(0, T; V^*)$ , formula (3.1.1) gives a bijective correspondence between initial states  $u_0$  in  $H$  and terminal states  $u(T)$  in  $y_f + D(e^{TA})$ .*

## 3.2. Well-posedness of the final value problem

In this section, we shall see that the bijection obtained in the previous section leads to well-posedness of the final value problem:

$$\begin{cases} u' + \mathcal{A}u = f, \\ u(T) = u_T. \end{cases} \quad (3.2.1)$$

First, using (3.1.15), we rewrite the bijection from Theorem 3.1.10 as

$$u_0 = e^{TA}(u_T - \int_0^T e^{-(T-s)\mathcal{A}} f(s) \, ds) = e^{TA}(u_T - y_f), \quad (3.2.2)$$

and note that since  $e^{TA} : D(e^{TA}) \rightarrow H$  allows substitution of (3.2.2), as initial state, into the representation formula in (2.4.2), the solution  $u$  to the final value problem is given by

$$u(t) = e^{-tA} e^{TA}(u_T - y_f) + \int_0^t e^{-(t-s)\mathcal{A}} f(s) \, ds. \quad (3.2.3)$$

Using this, we obtain the following theorem.

**Theorem 3.2.1.** *Let  $V \subseteq H \subseteq V^*$  algebraically, topologically and densely,  $A$  a Lax-Milgram operator defined from the triple  $(H, V, a(\cdot, \cdot))$ , where  $a(\cdot, \cdot)$  is a bounded and  $V$ -elliptic sesquilinear form, and with extension  $\mathcal{A} : V \rightarrow V^*$ . For given  $f \in L_2(0, T; V^*)$ , and  $u_T \in H$ , the condition*

$$u_T - y_f \in D(e^{TA}) \quad (3.2.4)$$

*is necessary and sufficient for the existence of some*

$$u \in X := L_2(0, T; V) \cap C([0, T]; H) \cap H^1(0, T; V^*) \quad (3.2.5)$$

*that solves (3.2.1).*

*In the affirmative case,  $u$  is uniquely determined, and has the representation in (3.2.3).*

For the space  $X$  of solutions in (3.2.5), we use the norm given by (cf. (2.2.27))

$$\|u\|_X^2 = \int_0^T \|u(t)\|_V^2 \, dt + \sup_{t \in [0, T]} \|u(t)\|^2 + \int_0^T \|u'(t)\|_{V^*}^2 \, dt. \quad (3.2.6)$$

*Proof.* When  $u_T$  and  $f$  fulfill the compatibility relation in (3.2.4), then  $u_0 = e^{TA}(u_T - y_f)$  is a well-defined vector in  $H$ , which is the unique initial state corresponding to  $u_T \in y_f + D(e^{TA})$ ; cf. Theorem 3.1.10 and (3.2.2). Moreover,



---

$u$  in (3.2.3) is the unique solution to the differential equation  $(\partial_t - \mathcal{A})u = f$  with initial state  $u_0$  (and final state  $u_T$ ). That  $u \in X$  follows from Theorem 2.2.1. Hence, (3.2.1) is uniquely solvable and the solution has the representation in (3.2.3).

On the other hand; if (3.2.1) has a solution  $u \in X$ , then  $u_T$  is reachable from  $u(0)$ , and it follows from the representation in (2.4.2) that

$$u_T = e^{-TA}u(0) + y_f. \quad (3.2.7)$$

In particular  $u_T - y_f \in D(e^{TA})$ . □

**Remark 3.2.2.** We see that not only is (3.2.1) solvable if and only if there exist a  $v \in H$  such that  $u_T - y_f = e^{-TA}v$ ; in the affirmative case, this  $v$  must also be  $u(0)$ . ○

**Remark 3.2.3.** Writing (3.2.4) as  $u_T = e^{-TA}u_0 + y_f$ , cf. Remark 3.2.2, we note that the conclusion of Theorem 3.2.1 is quite natural, since it follows that the admissible terminal data  $u_T$  when compared to the solution of

$$\begin{cases} u' + \mathcal{A}u = f, \\ u(0) = u_0, \end{cases} \quad (3.2.8)$$

is a sum of the reachable terminal state,  $e^{-TA}u_0$ , of the semi-homogeneous equation with  $f = 0$  and the reachable terminal state  $y_f$  of the semi-homogeneous equation with  $u_0 = 0$ . ○

**Remark 3.2.4.** To elaborate on the constraint  $u_T \in y_f + D(e^{TA})$ , we consider, on the solution space  $X$ , the following matrix formed operator

$$\begin{pmatrix} \partial_t + \mathcal{A} \\ r_T \end{pmatrix} : X \rightarrow \begin{matrix} L_2(0, T; V^*) \\ \times \\ H \end{matrix}. \quad (3.2.9)$$

Then (3.2.1) obviously has a solution if and only if

$$\begin{pmatrix} f \\ u_T \end{pmatrix} \in R\left(\begin{pmatrix} \partial_t + \mathcal{A} \\ r_T \end{pmatrix}\right), \quad (3.2.10)$$

To test whether the pair  $(f, u_T)$  belongs to the range in (3.2.10), one may use the operator

$$K : L_2(0, T; V^*) \times H \rightarrow H, \quad K(f, u_T) := u_T - \int_0^T e^{-(T-t)\mathcal{A}} f(t) dt.$$

---

Then Theorem 3.2.1 states that

$$R\left(\begin{pmatrix} \partial_t + \mathcal{A} \\ r_T \end{pmatrix}\right) = K^{-1}(D(e^{TA})). \quad (3.2.11)$$

◦

We now introduce the space of admissible data

$$Y := \{(f, u_T) \in L_2(0, T; V^*) \times H \mid u_T - y_f \in D(e^{TA})\}. \quad (3.2.12)$$

It is immediate to see that  $Y$  may be endowed with the obvious inner product

$$\begin{aligned} & (u_T \mid v_T) + (f \mid g)_{L_2(0, T; V^*)} + (u_T - y_f \mid v_T - y_g) \\ & + (e^{TA}(u_T - y_f) \mid e^{TA}(v_T - y_g)), \end{aligned} \quad (3.2.13)$$

and by Remark 3.2.4 we see that  $Y = K^{-1}(D(e^{TA}))$ . Moreover,  $Y$  is complete, and as for the initial value problem, we have continuous dependence on the data.

**Theorem 3.2.5.** *The solution  $u$  depends continuously on the data  $u_T$  and  $f$ , provided with the induced norm*

$$\|(f, u_T)\|_Y^2 := \|u_T\|_H^2 + \|f\|_{L_2(0, T; V^*)}^2 + \|u_T - y_f\|_{D(e^{TA})}^2. \quad (3.2.14)$$

Furthermore, the space of admissible data  $Y$  is a Hilbert space.

*Proof.* The continuity follows from Corollary 2.2.6 by inserting  $u_0 = e^{TA}(u_T - y_f)$ , cf. (3.2.2), in (2.2.28), yielding

$$\|u\|_X^2 \leq c(\|e^{TA}(u_T - y_f)\|^2 + \|f\|_{L_2(0, T; V^*)}^2) \quad (3.2.15)$$

$$\leq c(\|u_T - y_f\|_{D(e^{TA})}^2 + \|f\|_{L_2(0, T; V^*)}^2) \leq c\|(f, u_T)\|_Y^2. \quad (3.2.16)$$

To show completeness, let  $\{(f_n, u_T^n)\}$  be a Cauchy sequence in  $Y$ , which is a subset of  $L_2(0, T; V^*) \times H$  provided with the norm in (3.2.14). Then,  $\{f_n\}$  is Cauchy in  $L_2(0, T; V^*)$  with  $f_n \rightarrow f \in L_2(0, T; V^*)$ , and  $\{u_T^n\}$  is Cauchy in  $H$  with  $u_T^n \rightarrow u_T \in H$ , by the completeness of  $L_2(0, T; V^*)$ , and  $H$  respectively.

We shall show that  $(f, u_T) \in Y$ , i.e. that  $u_T - y_f \in D(e^{TA})$ , and that

$$\|u_T^n - y_{f_n} - (u_T - y_f)\|_{D(e^{TA})} \rightarrow 0. \quad (3.2.17)$$

Note that Lemma 3.1.1 implies that  $\{y_{f_n}\}$  converges to  $y_f$  in  $H$ . Hence,  $u_T^n - y_{f_n} \rightarrow u_T - y_f$  in  $H$  for  $n \rightarrow \infty$ .

Since  $u_T^n - y_{f_n} \in D(e^{TA})$ , there is a  $g_n \in H$  with  $g_n = e^{TA}(u_T^n - y_{f_n})$ . Now, Lemma 3.1.5 gives that

$$\|g_n\| \leq \|u_T^n - y_{f_n}\|_{D(e^{TA})}. \quad (3.2.18)$$

---

So  $\{g_n\}$  is Cauchy in  $H$ , since  $\{u_T^n - y_{f_n}\}$  is so in  $D(e^{TA})$ . Therefore, there is a  $g \in H$  such that  $\|g_n - g\| \rightarrow 0$ . Since  $e^{-TA} \in \mathbb{B}(H)$  it follows that, in  $H$ ,

$$u_T - y_f = \lim_{n \rightarrow \infty} (u_T^n - y_{f_n}) = \lim e^{-TA} g_n = e^{-TA} g. \quad (3.2.19)$$

Hence,  $u_T - y_f \in D(e^{TA})$ . Moreover,

$$\|e^{TA}(u_T^n - y_{f_n} - u_T - y_f)\| = \|g_n - g\| \rightarrow 0, \quad (3.2.20)$$

which yields (3.2.17) □

Now, by Theorem 3.2.1 and Theorem 3.2.5 together, we have the following theorem

**Theorem 3.2.6.** *For  $(f, u_T) \in Y$ , the final value problem*

$$\begin{cases} u' + \mathcal{A}u = f, \\ u(T) = u_T \end{cases} \quad (3.2.21)$$

*is well-posed with  $u \in X$ .*

### 3.3. Notes

The idea of using semigroup theory to study ill-posed parabolic problems is not new. It has e.g. been used in the method of quasi reversability to give approximative solutions to the parabolic final value problem by altering the operator, or, in the quasi boundary reversability method, the terminal data. The method of quasi reversability was introduced by Jacques-Louis Lions and Robert Lattès in [LL67], and the assumptions on the operator  $A$  was weakened e.g. by Ralph Showalter in [Sho74]. In fact, the result in [Sho74], which concerns the homogeneous equation, is related to the result in the present chapter. In [Sho74], the result is that, assuming that the operator is m-accretive with semiangle  $\pi/4$ , at most one solution exists; and furthermore, that a solution exists if and only if the final state can be approximated by the group generated by the Yosida approximation of the operator. As Lax-Milgram operators are generators of a semigroups after a change of sign, they are indeed m-accretive, but we work under the assumption that the operator is  $V$ -elliptic. However, our semiangle is  $\theta_0 = \cot(C_3 C_4^{-1})$ , and it is just in  $]0, \pi/2]$ , so our theory is more general in this respect. Furthermore, in the present chapter, we gave an explicit characterization of the space of admissible data, which fits well with the difficulties of solving the parabolic problem backward in time, described e.g. in [Isa98, Chapter 3] and [Pay75], and we have showed that a solution, even in the weak sense, does only exist for a very restricted class of terminal data.

---

We point out that it is expected to have some compatibility relations on the data in order to have solvability. Gerd Grubb and Vsevolod Solonikov gave in [GS90] explicit, local compatibility relations for the initial parabolic problem in a general setting to be solvable in certain Sobolev spaces. We note that our compatibility relation has a more implicit, and non-local nature because of the integral from zero to  $T$ , and the semigroup operators in all terms.

## 4. Applications to the heat equation

The purpose of this chapter is to show well-posedness of the final value heat equation by application of the results obtained in the previous chapters.

Consider therefore the final value heat equation

$$\begin{cases} (\partial_t - \Delta)u = f, & \text{in } Q := ]0, T[ \times \Omega \\ \gamma_0 u = g, & \text{on } \Gamma := ]0, T[ \times \partial\Omega \\ r_T u = u_T, & \text{at } \{T\} \times \Omega, \end{cases} \quad (4.0.1)$$

where  $\Omega$  is assumed to be bounded with  $\partial\Omega$  smooth (cf. Appendix A). With homogeneous boundary conditions, i.e. with  $g = 0$ , this is a special case of the theory treated in Section 3.2, cf. the description in Section 4.1 below. The case with inhomogeneous boundary data is treated in Section 4.2.

### 4.1. Homogeneous boundary data

When  $g \equiv 0$  in (4.0.1), let

$$V = H_0^1(\Omega), \quad H = L_2(\Omega) \quad \text{and} \quad V^* = H^{-1}(\Omega). \quad (4.1.1)$$

The Dirichlet realization of the Laplacian,  $-\Delta_{\gamma_0}$ , is the Lax-Milgram operator associated with the triple  $(L_2(\Omega), H_0^1(\Omega), s)$ , where

$$s(u, v) = \sum_{j=1}^n (\partial_j u \mid \partial_j v)_{L_2(\Omega)}. \quad (4.1.2)$$

We note that  $s(u, v)$  is  $H_0^1(\Omega)$ -elliptic and symmetric, so  $A = -\Delta_{\gamma_0}$  is selfadjoint on  $L_2(\Omega)$ . Also, we denote by  $-\mathcal{A} : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$  the extension of  $-\Delta_{\gamma_0}$ . In particular,  $-(-\Delta_{\gamma_0}) = \Delta_{\gamma_0}$  is the generator of the analytic semigroup on  $L_2(\Omega)$ , as is  $\mathcal{A}$  on  $H^{-1}(\Omega)$  by Lemma 2.3.3, and Lemma 2.3.5 respectively.

Consistent with the notation in general semigroup theory, we let  $e^{t\Delta_{\gamma_0}}$  denote the semigroup generated by  $\Delta_{\gamma_0}$ . Moreover, we denote the inverse of this semigroup by

$$(e^{t\Delta_{\gamma_0}})^{-1} = e^{-t\Delta_{\gamma_0}}, \quad (4.1.3)$$

although this notation requires some diligence as noted in Section 3.1.

Since  $H_0^1(\Omega)$  is compactly embedded in  $L_2(\Omega)$  by Rellich's Theorem (see e.g. [Gru09, Theorem 8.2]), we have by Remark 2.1.3, that  $L_2(\Omega)$  has an orthonormal basis of eigenvectors of  $-\Delta_{\gamma_0}$ . Let  $\{e_j\}$  denote this basis with the corresponding eigenvalues given by  $-\Delta_{\gamma_0} e_j = \lambda_j e_j$ , and note that  $0 < \lambda_j, \forall j$ , by the  $H_0^1$ -ellipticity, and that we may organize the eigenvalues such that they form an increasing sequence,  $\lambda_1 \leq \lambda_2 \leq \dots$  (cf. Remark 2.1.3). We will henceforth assume this ordering of the eigenvalues.

As an immediate consequence to Theorem 3.2.1, and Theorem 3.2.5, we have the following proposition:

**Proposition 4.1.1.** *Let  $-\Delta_{\gamma_0}$  be the Lax-Milgram operator associated to the triple*

*( $L_2(\Omega), H_0^1(\Omega), s(\cdot, \cdot)$ ), and let  $-\mathcal{A}$  denote the extension of  $-\Delta_{\gamma_0}$ , which maps  $H_0^1(\Omega)$  onto  $H^{-1}(\Omega)$ . Assume that  $f \in L_2(0, T; H^{-1}(\Omega))$ ,  $u_T \in L_2(\Omega)$ , and that*

$$u_T - \int_0^T e^{(T-s)\mathcal{A}} f(s) ds \in D(e^{-T\Delta_{\gamma_0}}). \quad (4.1.4)$$

*Then, the final value problem*

$$\begin{cases} (\partial_t - \Delta)u = f, & \text{in } Q \\ \gamma_0 u = 0, & \text{on } \Gamma \\ r_T u = u_T, & \text{at } \{T\} \times \Omega \end{cases} \quad (4.1.5)$$

*is well-posed, i.e. it has a uniquely determined solution*

$$u \in X = L_2(0, T; H_0^1(\Omega)) \cap C([0, T]; L_2(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)), \quad (4.1.6)$$

*which fulfils*

$$\|u\|_X^2 \leq c \|(f, u_T)\|_Y^2 = \|u_T\|_{L_2(\Omega)}^2 + \|f\|_{L_2(0, T; H^{-1}(\Omega))}^2 + \|u_T - y_f\|_{D(e^{TA})}^2. \quad (4.1.7)$$

*In the affirmative case,*

$$u_T - \int_0^T e^{(T-s)\mathcal{A}} f(s) ds = e^{T\Delta_{\gamma_0}} u(0). \quad (4.1.8)$$

*Moreover, it should be added, that the compatibility relation (4.1.4) is also necessary for the existence of a solution, cf. Theorem 3.2.1.*

We will now generalize to study the well-posedness in the case of inhomogeneous boundary data.

## 4.2. Inhomogeneous boundary data

As for the general parabolic evolution problem, the result regarding the solvability of the final value problem follows from solvability results regarding the initial value problem. Hence, we start by considering the initial value heat equation with inhomogeneous boundary data:

$$\begin{cases} (\partial_t - \Delta)u = f, & \text{in } Q := ]0, T[ \times \Omega \\ \gamma_0 u = g, & \text{on } \Gamma := ]0, T[ \times \partial\Omega \\ r_0 u = u_0, & \text{at } \{0\} \times \Omega. \end{cases} \quad (4.2.1)$$

We show results similar to those of Theorem 2.2.1, Corollary 2.2.6 and Theorem 2.4.1. The idea is to use the linearity of (4.2.1) to rewrite it to an equivalent problem with homogeneous boundary data with the aid of a function  $w$ , which satisfies that  $\gamma_0 w = g$  on the boundary  $\Gamma$ . Note that we use the notation  $\gamma_0$  to denote both the trace operator on  $Q$ , as well as that on  $\Omega$ . The meaning should however be clear from the context.

We have the following result regarding such a  $w$ :

**Lemma 4.2.1.** *For  $g \in H^{1/2,1/2}(\Gamma)$  there exist a  $w \in H^{1,1}(Q)$  such that  $\gamma_0 w = g$ . Furthermore, there exist a bounded right inverse  $\tilde{K}_0$  of  $\gamma_0$ , so with  $w = \tilde{K}_0 g$ ,  $w$  is uniquely determined from  $g$ , and*

$$\|w\|_{H^{1,1}(Q)} \leq c \|g\|_{H^{1/2,1/2}(\Gamma)}. \quad (4.2.2)$$

We use the notation  $H^{s,s}(Q)$  (cf. (A.0.19)) even though it is straightforward to show that it is equivalent to  $H^s(Q)$ , but we keep this notation in order to make the connection to [LM72a] more transparent.

*Proof.* By [LM72a, Proposition 4.2.1], we write  $H^{1/2,1/2}(\Gamma)$  as an interpolation space,

$$H^{1/2,1/2}(\Gamma) = [H^{1,1}(\Gamma), L_2(\Gamma)]_{1/2}. \quad (4.2.3)$$

Then, by Theorem 1.4.2 in [LM72b], using that  $\Omega$  locally looks like  $\mathbb{R}_+^n$  (cf. Appendix A), there exist  $w$  such that  $\hat{w} \in L_2(\mathbb{R}_{\xi_n}; H^{1,1}(\Gamma))$ , and  $|\xi_n| \hat{w} \in L_2(\mathbb{R}_{\xi_n}; L_2(\Gamma))$ , that fulfills  $\gamma_0 w = g$ . More precisely, there is a  $w \in H^{1,1}(Q)$  which is equal to  $g$  on  $\Gamma$ . In particular,  $\gamma_0 : H^{1,1}(Q) \rightarrow H^{1/2,1/2}(\Gamma)$  is surjective, and in fact bounded (see [LM72b, page 21-22]).

Now,  $\gamma_0 : H^{1,1}(Q) \ominus Z(\gamma_0) \rightarrow H^{1/2,1/2}(\Gamma)$  is bijective, and

$$\tilde{K}_0 : H^{1/2,1/2}(\Gamma) \rightarrow H^{1,1}(Q) \ominus Z(\gamma_0) \quad (4.2.4)$$

is the corresponding inverse. Since  $\gamma_0$  is bounded,  $Z(\gamma_0)$  is closed, hence complete in the norm on  $H^{1,1}(Q)$  restricted hereto. So by the open mapping principle (see e.g. [Lax02, page 170]),  $\tilde{K}_0$  is bounded and (4.2.2) follows.  $\square$

When considered as a mapping  $H^{1/2,1/2}(\Gamma) \rightarrow H^{1,1}(Q)$ , then  $\tilde{K}_0$  is a right inverse, i.e.  $\gamma_0 \tilde{K}_0 = I_{H^{1/2,1/2}(\Gamma)}$ . We shall henceforth let  $w := \tilde{K}_0$  be the particular choice of  $w$  introduced in the previous lemma. We have the following result:

**Proposition 4.2.2.** *The heat problem in (4.2.1) has a unique solution*

$$u \in X' := L_2(0, T; H^1(\Omega)) \cap C([0, T]; L_2(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)). \quad (4.2.5)$$

Furthermore, it depends continuously on the data, i.e.

$$\|u\|_{X'}^2 \leq c(\|u_0\|_{L_2(\Omega)}^2 + \|f\|_{L_2(0, T; H^{-1}(\Omega))}^2 + \|g\|_{H^{1/2,1/2}(\Gamma)}^2). \quad (4.2.6)$$

We note that the norm on  $X'$  is given by

$$\|x\|_{X'}^2 = \|x\|_{L_2(0, T; H^1(\Omega))}^2 + \sup_{0 \leq t \leq T} \|x(t)\|_{L_2(\Omega)}^2 + \|x\|_{H^1(0, T; H^{-1}(\Omega))}^2, \quad (4.2.7)$$

quite similar to (2.2.27), and turn to the proof of the proposition.

*Proof.* Taking  $w$  as in Lemma 4.2.1, we let  $v := u - w$  and note that (4.2.1) is equivalent to

$$\begin{cases} (\partial_t - \Delta)v = \tilde{f}, & \text{in } Q := ]0, T[ \times \Omega \\ \gamma_0 v = 0, & \text{on } \Gamma := ]0, T[ \times \partial\Omega \\ r_0 v = \tilde{u}_0, & \text{at } \{0\} \times \Omega \end{cases} \quad (4.2.8)$$

with

$$\tilde{f} := f - (\partial_t - \Delta)w \quad \text{and} \quad \tilde{u}_0 := u_0 - w(0). \quad (4.2.9)$$

Now,  $w \in H^{1,1}(Q)$  implies that  $w \in C([0, T]; L_2(\Omega))$ , by Lemma 2.2.2, so that  $w(0)$  is well defined. Since  $w \in H^1(0, T; L_2(\Omega)) \subseteq H^1(0, T; H^{-1}(\Omega))$ ,  $w$  is in  $X'$ . The inequality

$$\begin{aligned} \|\partial_s w\|_{L_2(0, T; H^{-1}(\Omega))}^2 + \|\Delta w\|_{L_2(0, T; H^{-1}(\Omega))}^2 &\leq C_1^2 \|w\|_{H^1(0, T; L_2(\Omega))}^2 + c \|w\|_{L_2(0, T; H^1(\Omega))}^2 \\ &\leq c \|w\|_{H^{1,1}(Q)}^2, \end{aligned} \quad (4.2.10)$$

implies that  $\partial_t w \in L_2(0, T; H^{-1}(\Omega))$ , and  $-\Delta w \in L_2(0, T; H^{-1}(\Omega))$ . It follows that  $\tilde{f} \in L_2(0, T; H^{-1}(\Omega))$ , and  $\tilde{u}_0 \in L_2(\Omega)$ , so by Theorem 2.2.1, the problem in (4.2.8) has a unique solution (cf. (2.2.26))

$$v \in X = L_2(0, T; H_0^1(\Omega)) \cap C([0, T]; L_2(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)). \quad (4.2.11)$$

Then, since  $w = \tilde{K}_0 g$  is uniquely determined from  $g$ , and  $w \in X'$ , the problem in (4.2.1) has a unique solution  $u = v + w \in X'$ .



We now show that the solution  $u$  depends continuously on the data, and note that by (2.2.8),

$$\sup_{0 \leq t \leq T} \|w(t)\|_{L_2(\Omega)} \leq c(\|w\|_{L_2(0,T;L_2(\Omega))} + \|\partial_s w\|_{L_2(0,T;L_2(\Omega))}), \quad (4.2.12)$$

which implies that

$$\|w\|_{X'}^2 \leq c\|w\|_{H^{1,1}(Q)}^2. \quad (4.2.13)$$

Now, by Corollary 2.2.6, (4.2.10) and (4.2.2),

$$\begin{aligned} \|u\|_{X'}^2 &\leq 2(\|v\|_X^2 + \|w\|_{X'}^2) \leq c(\|\tilde{u}_0\|_{L_2(\Omega)}^2 + \|\tilde{f}\|_{L_2(0,T;H^{-1}(\Omega))}^2 + \|w\|_{X'}^2) \\ &\leq c(\|u_0\|_{L_2(\Omega)}^2 + \|f\|_{L_2(0,T;H^{-1}(\Omega))}^2 + \|w(0)\|_{L_2(\Omega)}^2 \\ &\quad + \|(\partial_t - \Delta)w\|_{L_2(0,T;H^{-1})}^2 + \|w\|_{X'}^2) \\ &\leq c(\|u_0\|_{L_2(\Omega)}^2 + \|f\|_{L_2(0,T;H^{-1}(\Omega))}^2 + \|g\|_{H^{1/2,1/2}(\Gamma)}^2), \end{aligned} \quad (4.2.14)$$

which concludes the proof.  $\square$

We shall obtain a representation for the solution in Proposition 4.2.2, but before we do so, we must introduce some notions needed.

It becomes useful to also introduce a particular right inverse of  $\gamma_0$  on  $\Omega$ . For this purpose, we note that, by a straightforward argument, the bijectiveness of  $-\mathcal{A} = -\Delta|_{H_0^1(\Omega)}$  implies that  $H^1(\Omega)$  is the direct sum of the kernel of  $\gamma_0$ , and the harmonic functions, i.e.

$$H^1(\Omega) = H_0^1(\Omega) \dot{+} Z(-\Delta). \quad (4.2.15)$$

**Remark 4.2.3.** *In fact,  $H_0^1(\Omega)$  is orthogonal to  $Z(-\Delta)$  in  $H^1(\Omega)$  with respect to the inner product  $(u|v)_{H^1(\Omega)} = \sum_{j=1}^n (\partial_j u | \partial_j v)_{L_2(\Omega)}$ , which by the Poincaré inequality (see e.g. [Tem01]), induces a norm equivalent to the usual norm on  $H^1(\Omega)$ , whenever  $\Omega$  is bounded. So, for any  $u \in Z(-\Delta)$ ,  $\phi \in C_0^\infty(\Omega)$*

$$(u|v)_{H^1(\Omega)} = \sum_{j=1}^n \langle \partial_j u, \overline{\partial_j \phi} \rangle = \langle -\Delta u, \overline{\phi} \rangle = 0, \quad (4.2.16)$$

and hence, by a limit argument, it holds for all  $u \in Z(-\Delta)$ ,  $v \in H_0^1(\Omega)$ , i.e.

$$H^1(\Omega) = H_0^1(\Omega) \oplus Z(-\Delta). \quad (4.2.17)$$

◦

---

Let  $P := \mathcal{A}^{-1}\Delta$ . Then it is easy to see that  $P^2 = P$ , and  $P \in \mathbb{B}(H^1(\Omega), H_0^1(\Omega))$ , so in fact  $P$  is the projection on  $H_0^1(\Omega)$  along  $Z(-\Delta)$ . Let  $Q = I - P$  denote the projection on  $Z(-\Delta)$  along  $H_0^1(\Omega)$ . We keep this notation for the projection on  $Z(-\Delta)$  even though we also use the notation  $Q$  for the time cylinder. The meaning should however be clear from the context.

Since  $\gamma_0 : H^1(\Omega) \rightarrow H^{1/2}(\partial\Omega)$  is a surjection (see e.g. [LM72b, Theorem 1.9.4]), we may choose an inverse, denoted by  $K_0$ , on the complement  $Z(-\Delta)$  of  $Z(\gamma_0)$ , i.e.

$$K_0 : H^{1/2}(\partial\Omega) \rightarrow Z(-\Delta). \quad (4.2.18)$$

In fact, since  $Z(-\Delta)$  is complete in the norm on  $H^1(\Omega)$  restricted hereto, we get that  $K_0$  is bounded by the open mapping principle (see e.g. [Lax02, page 170]), analogue to the proof of Lemma 4.2.1. With this choice of  $K_0$ , we have

$$K_0\gamma_0 = K_0\gamma_0(P + Q) = K_0\gamma_0Q = I_{Z(-\Delta)}Q = Q. \quad (4.2.19)$$

As before, when we consider  $K_0$  as a map  $H^{1/2}(\partial\Omega) \rightarrow H^1(\Omega)$ , it is a right inverse for  $\gamma_0$ , i.e.  $\gamma_0 K_0 = I_{H^{1/2}(\partial\Omega)}$ .

This observation is used to prove the following lemma.

**Lemma 4.2.4.** *For  $w \in H^{1,1}(Q)$  with  $\gamma_0 w = g$ , the integral*

$$\int_0^{T-1/n} (-\Delta_{\gamma_0}) e^{(T-s)\Delta_{\gamma_0}} w(s) \, ds \quad (4.2.20)$$

*converges in  $L_2(\Omega)$ , and the integral*

$$\int_0^{T-1/n} \mathcal{A} e^{(T-s)\mathcal{A}} K_0 g(s) \, ds \quad (4.2.21)$$

*converges in  $H^{-1}(\Omega)$ .*

*Proof.* The strong continuity of the semigroup (cf. Definition 2.3.1), and the uniform boundedness principle yields that  $\|e^{(T-(T-1/n))\Delta_{\gamma_0}}\|_{\mathbb{B}(L_2(\Omega))} \leq c$ , and using bilinearity we obtain

$$e^{(T-(T-1/n))\Delta_{\gamma_0}} w(T-1/n) \rightarrow w(T) \text{ in } L_2(\Omega). \quad (4.2.22)$$

By Proposition 2.3.9,  $\partial_s(e^{(T-s)\Delta_{\gamma_0}} w(s)) = -\Delta_{\gamma_0} e^{(T-s)\Delta_{\gamma_0}} w(s) + e^{(T-s)\Delta_{\gamma_0}} \partial_s w(s)$ , where both terms are in  $L_1(0, T-1/n; L_2(\Omega))$  for all  $n$ , so by Lemma 2.2.2,

$$\begin{aligned} [e^{(T-s)\Delta_{\gamma_0}} w(s)]_{s=0}^{s=T-1/n} &= \int_0^{T-1/n} (-\Delta_{\gamma_0}) e^{(T-s)\Delta_{\gamma_0}} w(s) \, ds \\ &\quad + \int_0^{T-1/n} e^{(T-s)\Delta_{\gamma_0}} \partial_s w(s) \, ds. \end{aligned} \quad (4.2.23)$$

Since the lefthand side, and the rightmost term both converge in  $L_2(\Omega)$  as  $n \rightarrow \infty$ , the integral  $\int_0^{T-1/n} (-\Delta_{\gamma_0})e^{(T-s)\Delta_{\gamma_0}}w(s) ds$  converges in  $L_2(\Omega)$  as claimed.

Naturally, this convergence also holds in  $H^{-1}(\Omega)$  by the continuous embeddings in (2.1.2), and since (4.2.19) gives  $Qw(s) = K_0\gamma_0w(s) = K_0g(s)$ ,

$$\begin{aligned} & \int_0^{T-1/n} \Delta_{\gamma_0} e^{(T-s)\Delta_{\gamma_0}} w(s) ds = \\ & \int_0^{T-1/n} [\mathcal{A}e^{(T-s)\mathcal{A}}Qw(s) ds - e^{(T-s)\mathcal{A}}(-\mathcal{A})Pw(s)] ds = \\ & \int_0^{T-1/n} [\mathcal{A}e^{(T-s)\mathcal{A}}K_0g(s) ds - e^{(T-s)\mathcal{A}}(-\mathcal{A})Pw(s)] ds. \end{aligned} \quad (4.2.24)$$

In the last term,  $e^{(T-s)\mathcal{A}}\mathcal{A}Pw(s) = e^{(T-s)\mathcal{A}}\Delta Pw(s) \in L_2(0, T; H^{-1}(\Omega))$ , which implies the convergence of (4.2.21) in  $H^{-1}(\Omega)$ .  $\square$

We use the following notation for the limits in the above Lemma:

$$\oint_0^T (-\Delta_{\gamma_0})e^{(T-s)\Delta_{\gamma_0}}w(s) ds := \lim_{n \rightarrow \infty} \int_0^{T-1/n} (-\Delta_{\gamma_0})e^{(T-s)\Delta_{\gamma_0}}w(s) ds \quad \text{in } L_2(\Omega), \quad (4.2.25)$$

and

$$\oint_0^T \mathcal{A}e^{(T-s)\mathcal{A}}K_0g(s) ds := \lim_{n \rightarrow \infty} \int_0^{T-1/n} \mathcal{A}e^{(T-s)\mathcal{A}}K_0g(s) ds \quad \text{in } H^{-1}(\Omega). \quad (4.2.26)$$

Using this notation, we have the following result

**Proposition 4.2.5.** *Let  $u$  be the unique solution to (4.2.1) (cf. Proposition 4.2.2). Then  $u$  has the following representation*

$$u(t) = e^{t\Delta_{\gamma_0}}u_0 + \int_0^t e^{(t-s)\mathcal{A}}f(s) ds - \int_0^t \mathcal{A}e^{(t-s)\mathcal{A}}K_0g(s) ds. \quad (4.2.27)$$

*Proof.* By Theorem 2.4.1,  $v$  has the representation

$$\begin{aligned} v(t) &= e^{t\Delta_{\gamma_0}}\tilde{u}_0 + \int_0^t e^{(t-s)\mathcal{A}}\tilde{f}(s) ds \\ &= e^{t\Delta_{\gamma_0}}u_0 - e^{t\Delta_{\gamma_0}}w(0) + \int_0^t e^{(t-s)\mathcal{A}}f(s) ds - \int_0^t e^{(t-s)\mathcal{A}}(\partial_s - \Delta)w(s) ds. \end{aligned} \quad (4.2.28)$$

---

Rewriting the last term, using (4.2.23) and (4.2.24) yields

$$\begin{aligned}
\int_0^t e^{(t-s)\mathcal{A}}(\partial_s - \Delta)w(s) \, ds &= \int_0^t e^{(t-s)\mathcal{A}}\partial_s w(s) \, ds + \int_0^t e^{(t-s)\mathcal{A}}(-\Delta)Pw(s) \, ds \\
&= [e^{(t-s)\Delta_{\gamma_0}}w(s)]_{s=0}^{s=t} - \int_0^t (-\Delta_{\gamma_0})e^{(t-s)\Delta_{\gamma_0}}w(s) \, ds \\
&\quad + \int_0^t e^{(t-s)\mathcal{A}}(-\Delta)Pw(s) \, ds \\
&= w(t) - e^{t\Delta_{\gamma_0}}w(0) + \int_0^t \mathcal{A}e^{(t-s)\mathcal{A}}K_0g(s) \, ds.
\end{aligned} \tag{4.2.29}$$

Inserting this into (4.2.28), yields (4.2.27) since  $u = v + w$ .  $\square$

Next, we address the solvability of the final value problem

$$\begin{cases} (\partial_t - \Delta)u = f, & \text{in } Q := ]0, T[ \times \Omega, \\ \gamma_0 u = g, & \text{on } \Gamma := ]0, T[ \times \partial\Omega, \\ r_T u = u_T, & \text{at } \{T\} \times \Omega. \end{cases} \tag{4.2.30}$$

As before, the aim is to give conditions on the data  $f$ ,  $g$  and  $u_T$  such that the system is well-posed and as before, this had to do with the representation of the solution to the initial value problem. However, we note that the term  $\int_0^t \mathcal{A}e^{(t-s)\mathcal{A}}K_0g(s) \, ds$  is difficult to handle, so we briefly explore this term, and introduce the following notation

$$z_g := \int_0^T \mathcal{A}e^{(t-s)\mathcal{A}}K_0g(s) \, ds. \tag{4.2.31}$$

Note that  $g \mapsto z_g$  is a well defined linear map. The following result shows that it maps  $H^{1/2, 1/2}(\Gamma)$  into  $L_2(\Omega)$ .

**Lemma 4.2.6.** *When  $g \in H^{1/2, 1/2}(\Gamma)$ , then  $\|z_g\|_{L_2(\Omega)} \leq c\|g\|_{H^{1/2, 1/2}(\Gamma)}$ .*

*Proof.* By the representation formula in (4.2.27), with  $t = T$ , we see that  $z_g = u(T)$  for homogeneous data  $f = 0$ , and  $u_0 = 0$ . Proposition 4.2.2 now yields (cf. (4.2.6))

$$\|z_g\|_{L_2(\Omega)}^2 \leq \sup_{0 \leq t \leq T} \|u(t)\|_{L_2(\Omega)}^2 \leq c\|g\|_{H^{1/2, 1/2}(\Gamma)}, \tag{4.2.32}$$

which concludes the proof.  $\square$

---

We now present the solvability result for the final value heat equation with inhomogeneous boundary data.

**Theorem 4.2.7.** *Assume that  $g \in H^{1/2,1/2}(\Gamma)$ ,  $f \in L_2(0, T; H^{-1}(\Omega))$ ,  $u_T \in L_2(\Omega)$ . Then the compatibility relation:*

$$u_T - y_f + z_g \in D(e^{-T\Delta_{\gamma_0}}), \quad (4.2.33)$$

*is necessary and sufficient for the existence of some*

$$u \in X' := L_2(0, T; H^1(\Omega)) \cap C([0, T]; L_2(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)), \quad (4.2.34)$$

*that solves (4.2.30). In the affirmative case,  $u$  is uniquely determined, and have the representation*

$$u(t) = e^{t\Delta_{\gamma_0}} e^{-T\Delta_{\gamma_0}} (u_T - y_f + z_g) + \int_0^t e^{(t-s)\mathcal{A}} f(s) ds - \int_0^t \mathcal{A} e^{(t-s)\mathcal{A}} K_0 g(s) ds. \quad (4.2.35)$$

*Proof.* Only minor modifications is needed to the proof of Theorem 3.2.1, by noting that  $z_g \in L_2(\Omega)$ , so that the representation formula in (4.2.27) is a bijective correspondence between initial and terminal data.  $\square$

**Remark 4.2.8.** *As in Proposition 4.1.1, we can make the statement in the above theorem more precise: If a solution to (4.2.30) exists, not only must  $u_T - y_f + z_g \in D(e^{-T\Delta_{\gamma_0}})$ , but in the affirmative case we have that*

$$u_T - y_f + z_g = e^{T\Delta_{\gamma_0}} u(0). \quad (4.2.36)$$

$\circ$

Denoting the space of admissible data by

$$Y' := \left\{ (f, g, u_T) \in L_2(0, T; H^{-1}(\Omega)) \times H^{1/2,1/2}(\Gamma) \times L_2(\Omega) \mid \right. \quad (4.2.37)$$

$$\left. u_T - y_f + z_g \in D(e^{-T\Delta_{\gamma_0}}) \right\}, \quad (4.2.38)$$

equipped with the norm

$$\begin{aligned} \|(f, g, u_T)\|_{Y'}^2 &:= \|u_T\|_{L_2(\Omega)}^2 + \|g\|_{H^{1/2}(Q)}^2 + \|f\|_{L_2(0, T; H^{-1}(\Omega))}^2 \\ &\quad + \|u_T - y_f + \int_0^T \mathcal{A} e^{(T-s)\mathcal{A}} K_0 g(s) ds\|_{D(e^{-T\Delta_{\gamma_0}})}^2, \end{aligned} \quad (4.2.39)$$

and recalling that

$$\|x\|_{X'}^2 := \|x\|_{L_2(0, T; H^1(\Omega))}^2 + \sup_{0 \leq t \leq T} \|x(t)\|_{L_2(\Omega)}^2 + \|x\|_{H^1(0, T; H^{-1}(\Omega))}^2 \quad (4.2.40)$$

is the norm on the solution space  $X'$  (cf. (4.2.34)), we have the following result on the well-posedness of the fully inhomogeneous problem in (4.2.1):

---

**Corollary 4.2.9.** *The solution  $u$  in  $X'$  depends continuously on the data  $(f, g, u_T)$  in  $Y'$  provided with the norm in (4.2.39). Furthermore,  $Y'$  is complete in the norm  $\|\cdot\|_{Y'}$ .*

*Proof.* Continuity follows by inserting  $u_0 = e^{-T\Delta_{\gamma_0}}(u_T - y_f + z_g)$  from (4.2.36) into (4.2.6). The completeness of  $Y'$  is deduced from the boundedness of  $z_g$  from Lemma 4.2.6 with minor modifications of the proof of Theorem 3.2.5.  $\square$

### 4.3. Notes

We note that we could have taken a different approach to get a solvability result of the final value heat equation in (4.2.30). Namely, the compatibility relation in (4.2.33) would, as one would indeed expect, follow directly from the compatibility relation deduced in (4.1.4) by similar observations as those presented here. However, by taking the present approach, we gain in addition, the representation formulas in (4.2.27) and (4.2.35) by straightforward computations.

The approach of rewriting the problem with inhomogeneous boundary data via the trace  $\gamma_0$  is well-known, and is e.g. used in combination with semigroup theory by Donald Washburn in [Was76] to introduce a generalized solution to the semihomogeneous initial value problem with non-zero boundary data, that is similar to  $z_g$  in the present chapter. Furthermore, he addressed the difficulty of obtaining integrability results for the term where the generator acts on the semigroup under some high regularity assumptions on the boundary data, and it is in this setting his generalized solution is to be seen. However, we only need the boundary data to be in  $H^{1/2,1/2}(\Gamma)$  for the semihomogeneous initial value problem to be solvable.

The idea in the present thesis has been to reduce the data space sufficiently to deduce well-posedness. This was also the strategy by Willard Miranker in [Mir61] to show existence of a solution to the homogeneous, one-dimensional heat equation on  $\mathbb{R}_t \times \mathbb{R}_x$ . In [Mir61], the result is that for terminal data with compactly supported Fourier transform, the heat problem is solvable, and the corresponding initial state has again compactly supported Fourier transform. Furthermore, he gave a counterexample to well-posedness in the one-dimensional case when one does not restrict the final values sufficiently.

However, in the present thesis we showed that not alone does a solution to the fully inhomogeneous parabolic final value problem exist, under appropriate assumptions on the final states and the inhomogeneities; it is even well-posed, when studied on appropriate spaces. To our present knowledge, there has been no previous results in the literature regarding the well-posedness of the final value parabolic problem.

## 5. A few examples

In this chapter, we give a short presentation of some immediate applications of the results, and methods obtained in the previous chapters. In particular, we see that the solution to two semihomogenous problems for the same operator can have very different nature, and we give a short proof of approximate controllability of the heat equation with an inner control exerted on the entire domain.

### 5.1. A study of the two semihomogeneous problems

For Hilbert spaces  $V$  and  $H$ , let  $V \subseteq H$  algebraically, topologically and densely with  $V \hookrightarrow H$  continuous. Let  $a(\cdot, \cdot)$  be a bounded,  $V$ -elliptic sesquilinear form, and  $A$  the Lax-Milgram operator associated with the triple  $(H, V, a(\cdot, \cdot))$  with extension  $\mathcal{A} : V \rightarrow V^*$ .

Let us consider the following semi-homogeneous problem

$$\begin{cases} u' + \mathcal{A}u = f & \text{in } \mathcal{D}'(0, T; V^*), \\ u(0) = 0 & \text{in } H. \end{cases} \quad (5.1.1)$$

By Theorem 2.4.1, the solution at time  $t = T$ ,  $u(T)$  is  $y_f$  (cf. (3.1.2)) at least for the heat equation. Now, Proposition 3.1.2 yields that  $V \subseteq R(y_f)$ . That is, all of  $V$  may be reached at time  $T$  for some solution to (5.1.1). We therefore, in many cases, have some flexibility in this solution.

This is in contrast to the solution of other semi-homogeneous problem which we now explore; at least for normal Lax- Milgram operators. Hence, we assume that  $A$  is normal, and consider

$$\begin{cases} u' + Au = 0 & \text{in } \mathcal{D}'(0, T; H), \\ u(0) = u_0 & \text{in } H, \end{cases} \quad (5.1.2)$$

with solution  $u(t) = e^{-At}u_0$ , where  $e^{-tA}$  is an analytic semigroup.

We now introduce the following auxillary function

$$h(t) := \|u(t)\|^2, \quad t \in [0, T] \quad (5.1.3)$$

and note that  $h$  is continuous for  $t \geq 0$  and differentiable for  $t > 0$ .

We have the following Proposition

---

**Proposition 5.1.1.** *The function  $h$  introduced in (5.1.3) is strictly convex with respect to time. Furthermore, its growth is bounded at  $t = 0$  by  $-m(A) < 0$  (cf. (A.0.9)).*

This makes the solution to the semi-homogeneous problem in (5.1.2) is very rigid; it is very hard to chance the solution locally in time without violating the strict convexity.

*Proof.* Assume that  $u_0 \in D(A)$  with  $\|u_0\| = 1$ . By Lemma 2.1.1, and the mean value theorem, there is a  $\tau > 0$  such that

$$t^{-1}(h(t) - h(0)) = h'(\tau) = 2 \operatorname{Re}(-Au(\tau) \mid u(\tau)) = -2 \operatorname{Re}(e^{-\tau A} Au_0 \mid e^{-A\tau} u_0). \quad (5.1.4)$$

Taking the limit as  $t \rightarrow 0$ , and consequently as  $\tau \rightarrow 0$ , we see that

$$h'(0) = -2 \operatorname{Re}(Au_0 \mid u_0). \quad (5.1.5)$$

For  $t > 0$ ,  $u(t) \in D(A) = D(A^*)$ , since  $A$  is normal, and  $Au(t) \in D(A)$ , since  $e^{-tA}$  is an analytic semigroup. Hence,  $A^*Au(t)$ , and consequently  $AA^*u(t)$ , is defined, and we have that

$$\begin{aligned} h''(t) &= \frac{d}{dt}((u'(t) \mid u(t)) + (u(t) \mid u'(t))) \\ &= (A^2u(t) \mid u(t)) + 2(u'(t) \mid u'(t)) + (u(t) \mid A^2u(t)) \\ &= (Au(t) \mid A^*u(t)) + 2(Au(t) \mid Au(t)) + (A^*u(t) \mid Au(t)) \\ &= \|(A + A^*)u(t)\|^2 + ((A^*A - AA^*)u(t) \mid u(t)). \end{aligned} \quad (5.1.6)$$

Hence,  $h'' \geq 0$  if  $(A^*A - AA^*) \geq 0$ . This is obviously so since  $A$  is normal, so  $h$  is convex.

The strict convexity follows since  $m(A + A^*) = 2m(A) > 0$ , so the operator  $A + A^*$  is injective, which by the injectivity of  $e^{-tA}$  (cf. Proposition 3.1.3), implies that  $h''(t) > 0$ .

Furthermore, since  $\|u_0\| = 1$  by assumption,  $\|u(t)\|$  is differentiable, and

$$\frac{d}{dt}\|u(t)\| \Big|_{t=0} = \frac{1}{2}\|u(t)\|^{-1}h'(t) \Big|_{t=0} = -\frac{2 \operatorname{Re}(Au(t) \mid u(t))}{2\|u(t)\|} \Big|_{t=0} = -\operatorname{Re}(Au_0 \mid u_0). \quad (5.1.7)$$

Therefore,

$$\frac{d}{dt}\|u(t)\| \Big|_{t=0} \in \operatorname{Re} \nu(-A), \quad (5.1.8)$$



---

So,

$$\frac{d}{dt} \|u(t)\| \Big|_{t=0} \leq \sup \operatorname{Re}\{(-Ay|y) \mid y \in \mathbb{C}^n, \|y\| = 1\} = -m(A), \quad (5.1.9)$$

i.e. the growth of the norm of the solution is bounded at time zero by  $m(-A)$ .  $\square$

**Remark 5.1.2.** *We note that it suffices for the convexity argument, that  $A^*A - AA^* \geq 0$ , so it may suffice for  $A$  to be hyponormal. However, we are not familiar with criteria herefor.*  $\circ$

## 5.2. The heat equation with inner control

Consider the following control problem for the heat equation with an inner control function  $f$  acting on the control area  $\omega \subseteq \Omega$ :

$$\begin{cases} u' - \Delta u = f \mathbb{1}_\omega & \text{in } Q \\ \gamma_0 u = 0, & \text{on } \Gamma \\ u(0, x) = u_0 & \text{at } \{0\} \times \Omega. \end{cases} \quad (5.2.1)$$

The question under consideration in control theory, is whether we can choose a control function such that the solution  $u$  exists, and exhibits a particular behaviour at time  $T > 0$ .

Let  $E(T, u_0)$  denote the states reachable from the initial value  $u_0$  at time  $T$ , i.e. the  $u(T)$ 's for which there is an  $f$  such that  $u$  solves (5.2.1). As in [MZ04], we say that (5.2.1) is:

- Exactly controllable if  $E(T, u_0) = L_2(\Omega)$  for all initial states  $u_0 \in L_2(\Omega)$ .
- Approximately controllable if for every  $u_0 \in L_2(\Omega)$ ,  $E(T, u_0) \subseteq L_2(\Omega)$  densely, i.e. if for every  $v \in L_2(\Omega)$  and  $\varepsilon > 0$ , there exist an  $N > 0$ , and a sequence of control functions  $\{f_n\}$ , with associated solutions  $\{u_n\}$ , such that  $\|v - u_n\| < \varepsilon$ ,  $n \geq N$ .
- Null-controllable if  $0 \in E(T, u_0)$  for every  $u_0 \in L_2(\Omega)$ .

It is well-known that (5.2.1) cannot be exactly controllable if the area of control  $\omega \subseteq \Omega$  is a strict subset, due to the hypoellipticity of the heat equation. The following proposition shows that not only is  $\omega = \Omega$  a necessary condition for exact controllability; for final states  $v \in H_0^1(\Omega)$ , it is also a sufficient condition.

**Proposition 5.2.1.** *Given  $u_0 \in L_2(\Omega)$ , the control problem in (5.2.1) with  $\omega = \Omega$  is exactly controllable for  $v \in H_0^1(\Omega)$ .*

---

*Proof.* Let  $v$  be an arbitrary vector in  $H_0^1(\Omega)$ . For a given  $u_0 \in L_2(\Omega)$ , the translation  $v - e^{T\Delta_{\gamma_0}}u_0$  is in  $H_0^1(\Omega)$ , since  $D(\Delta_{\gamma_0}) = H^2(\Omega) \cap H_0^1(\Omega)$  (see e.g. [Gru09]). By Proposition 3.1.2,  $H_0^1(\Omega) \subseteq R(y_f)$ . Hence, there exists  $f \in L_2(0, T; H^{-1}(\Omega))$  such that

$$v - e^{T\Delta_{\gamma_0}}u_0 = y_f. \quad (5.2.2)$$

Therefore,  $v$  and  $f$  fulfills the compatibility relation  $v - y_f \in D(e^{-T\Delta_{\gamma_0}})$  and by Proposition 4.1.1,

$$\begin{cases} u' - \Delta u = f & \text{in } Q \\ \gamma_0 u = 0 & \text{on } \Gamma \\ u(T) = v & \text{at } \{T\} \times \Omega \end{cases} \quad (5.2.3)$$

is well-posed. Because of (5.2.2), we see that  $f$  is a control function steering  $u_0$  to  $v$ .  $\square$

**Remark 5.2.2.** *First, we note that the connection between  $u_0$  and  $v$  is very explicit, cf. (5.2.2). Furthermore, the proof is constructive, since we choose the control function  $f$  appropriately in the proof of Proposition 3.1.2. With the choice of  $f$  from this Proposition, the control function is constant in time.*  $\circ$

Furthermore, we have the following

**Proposition 5.2.3.** *Given  $u_0 \in L_2(\Omega)$ , the control problem in (5.2.1) with  $\omega = \Omega$  is approximately controllable for  $v \in L_2(\Omega)$ .*

*Proof.* By Remark 3.1.6,  $R(y_f) \subseteq L_2(\Omega)$  is dense, so for  $v \in L_2(\Omega)$ ,  $v - e^{TA}u_0$  is in  $L_2(\Omega)$ , and there exist a sequence  $\{w_n\} \in R(y_f)$  such that  $\|v - e^{TA}u_0 - w_n\|_{L_2(\Omega)} \rightarrow 0$ ,  $n \rightarrow \infty$ . Hence, there is a sequence of control functions  $\{f_n\} \in L_2(0, T; H^{-1}(\Omega))$  with  $w_n := y_{f_n}$ . Now,  $e^{TA}u_0 + y_{f_n} := u_T^n$  is a reachable state at time  $T$  for the system (5.2.1), and  $\|v - u_T^n\|_{L_2(\Omega)} \rightarrow 0$ . Whence, (5.2.1) is approximately controllable.  $\square$

## 5.3. Notes

Control theory is in itself a vast subject, and many contributions to that field have been made over the years. For the interested reader, we mention the surveys by Sorin Micu and Enrique Zuazua in [MZ04] and that of Enrique Fernández-Cara and Sergio Guerrero in [FCG06]. In both cases, approximate controllability of the heat equation, with an inner control function acting on  $\omega \subset \Omega$ , is proven with a variational approach. Namely, the approximate control is obtained as the minimizer of a particular functional, the existence which is guaranteed whenever this functional is coercive. This is accomplished using the Holmgren Uniqueness Theorem in [MZ04] and via Carleman-inequalities, used to proof the observability inequality, in [FCG06].

---

A different approach to the controllability of the heat equation was taken by David Russel in [Rus73]. Here a controllability result for the heat equation was presented, where the terminal data is taken in the domain for a particular unbounded operator, and with a control function acting on all of the boundary of the time cylinder. It was based on a similar control problem for the wave equation via a harmonic analysis approach.

## Conclusion and outlook

We have seen that the well-posedness of a problem is highly dependent on the spaces under consideration. Even though the final value parabolic evolution problem is generally considered to be ill-posed, we have succeeded in carefully matching the data space with the standard solution space in such a way that the problem becomes in fact well-posed. We furthermore provided a full characterization of the data space, which turns out to be a Hilbert space.

These general conclusions were then applied to the special case of the final value heat equation with homogeneous boundary data. In addition, we showed that similar observations apply to the heat equation with inhomogeneous boundary data, which is also well-posed. Again, we obtained a characterization of the space of admissible data.

Lastly, we showed some applications of the previous results, including a short proof of approximate controllability of the heat equation with an inner control.

Regarding future work, it could be interesting to explore the possibility of weakening the assumptions on the operator; is it e.g. possible to make similar conclusions if  $V$ -ellipticity is replaced by coercitivity.

In the field of control theory, it would be interesting to investigate if further conclusions could be made based on the approach presented here; it would be particularly interesting to explore whether it would be fruitful for the question of boundary controllability.

By comparison, since the Hilbert Uniqueness Method, applied to the wave equation in e.g. [Ped00], depends on well-posedness of the final value problem (here the hyperbolic one), it would be interesting to study the possibility of applying it to the heat equation.

# A. Notions and notation

The purpose of this chapter is to introduce some important notation, which will be used throughout the text without any further explanation. We also introduce some spaces which the text will depend heavily on, without stating all the details, but providing references when needed.

## General notation

As suggested by the title, this section provides some overall notation.

- Special constants will be denoted by  $C_i$ ,  $i \in \mathbb{N}$ , the value of which will refer to the equation in which the constant was introduced. Other constants will be denoted by  $c$ , and the value of these vary with the context.
- The ball in  $\mathbb{R}^n$  centered at  $x$  with radius  $r$  is denoted  $B(x, r)$ .
- The positive half-space of  $\mathbb{R}^n$  is denoted  $\mathbb{R}_+^n := \{(x', x_n) \mid x_n > 0\}$ , where  $x' \in \mathbb{R}^{n-1}$ .
- The transpose of a vector  $(a, b)$  is denoted  $(a, b)^T$ , so that

$$(a, b)^T = \begin{pmatrix} a \\ b \end{pmatrix}. \quad (\text{A.0.1})$$

- The laplacian on  $\mathbb{R}^n$  is denoted  $\Delta = \sum_{j=1}^n \partial_j^2$ .
- Trace operators are denoted  $\gamma_j$ . The reader is referred to [Gru09] or [LM72b].
- The Fourier transform  $\mathcal{S}'(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$  is denoted by

$$(\mathcal{F}_x u)(\xi) = \int_{\mathbb{R}^n} e^{-ix \cdot \xi} dx, \quad (\text{A.0.2})$$

or in short  $\hat{u}$ .

## Domains

In general, for a domain  $D \subset \mathbb{R}^n$ , we let  $D^\circ$  denote the interior,  $\partial D$  the boundary and  $\overline{D}$  the closure.

Now, we present the prerequisites to be fulfilled by the domain on which we will study the heat equation. We give a useful way to look at this domain, namely via local maps to the Euclidian space.

We let  $\Omega$  denote an open and bounded domain of  $\mathbb{R}^n$  which is smooth in the sense of Definition C.1 in [Gru09, Appendix C]. That is, we assume that  $\Omega$  is locally on one side of the boundary, which is smooth, so that at any boundary point, the outward normal can be taken to be in the positive  $x_n$ -direction after a diffeomorphism. Specifically, at every boundary point  $x$  with neighbourhood  $U$ , we require the existence of a  $C^\infty$ -diffeomorphism  $\kappa$  such that

$$\kappa(x) = 0 \tag{A.0.3}$$

$$\kappa(U \cap \Omega) = B(0, 1) \cap \mathbb{R}_+^n \tag{A.0.4}$$

$$\kappa(U \cap \partial\Omega) = B(0, 1) \cap \mathbb{R}^{n-1}. \tag{A.0.5}$$

We also introduce the time cylinder  $Q := ]0, T[ \times \Omega$  and its boundary  $\Gamma := ]0, T[ \times \partial\Omega$ .

## Banach spaces

For a Banach space  $X$ , we denote by  $X^*$  the anti-dual of  $X$ , i.e.  $X^*$  consists of the bounded conjugate linear functionals on  $X$ . Also,  $\langle \cdot, \cdot \rangle_{X^*, X}$  denotes the scalar product between  $X^*$  and  $X$ , and  $\| \cdot \|_X$  denotes the norm on  $X$ . If  $X$  is a Hilbert space, then  $(\cdot | \cdot)_X$  denotes the inner product on  $X$  which induces  $\| \cdot \|_X$ . In the special case of  $X = \mathbb{C}$ , the norm is denoted by  $|\cdot|$ .

For Banach spaces  $X, Y$ , we denote the norm of the pair  $(x, y)$  in the product space  $X \times Y$  by

$$\|(x, y)\|_{X \times Y}^2 = \|x\|_X^2 + \|y\|_Y^2, \tag{A.0.6}$$

and the norm on the intersection  $X \cap Y$  by

$$\|v\|_{X \cap Y}^2 = \|v\|_X^2 + \|v\|_Y^2. \tag{A.0.7}$$

We consider a mapping  $T : X \rightarrow Y$ , with  $X$  and  $Y$  Banach spaces. Then, we denote by  $R(T) := \{y \in Y \mid \exists x \in X : y = Tx\}$  the range of  $T$  and  $D(T)$  the domain of  $T$ , i.e. the subspace of  $X$  on which  $T$  is defined. We provide  $D(T)$  with the graph-norm, which we denote by  $\| \cdot \|_{D(T)}$ , i.e.

$$\|x\|_{D(T)}^2 = \|x\|_X^2 + \|Tx\|_Y^2. \tag{A.0.8}$$

---

Moreover,  $Z(T) := \{x \in D(T) \mid Tx = 0\}$  denotes the Null-space of  $T$  and  $T^{-1}(U)$  the pre-image of  $U \subseteq Y$ . If  $T$  is injective,  $T^{-1} : D(T^{-1}) \rightarrow X$  denotes the inverse of  $T$ .

We denote the set of bounded mappings  $X \rightarrow Y$  by  $\mathbb{B}(X, Y)$ .

When  $X$  is a Hilbert space and  $T : X \rightarrow X$ , the lower bound of  $T$  is given by

$$m(T) := \inf \{ \operatorname{Re} (Tx \mid x)_X \mid x \in D(T), \|x\|_X = 1 \}, \quad (\text{A.0.9})$$

and the numerical range of  $T$  by

$$\nu(A) := \{ (Tx \mid x)_X \mid x \in D(T), \|x\|_X = 1 \}. \quad (\text{A.0.10})$$

In this case, we also denote by  $\rho(T)$  the resolvent set of  $T$ , i.e. the set consisting of  $\lambda \in \mathbb{C}$  where  $(T - \lambda I)^{-1}$  is densely defined and bounded. Furthermore, we denote the spectrum of  $T$  by  $\sigma(T) := \mathbb{C} \setminus \rho(T)$ .

Of special interest will be the Hilbert spaces

$$H_0^1(\Omega) = \{u \in H^1(\Omega) \mid \gamma_0 = 0\} = \{u \in L_2(\Omega) \mid \partial u \in L_2(\Omega), \gamma_0 = 0\} \quad (\text{A.0.11})$$

and its dual  $H^{-1}(\Omega)$ .

## Vector valued functions

For a Banach space  $X$ , we define measurability of vector valued functions as in [RS80], and say that  $v : [0, T] \rightarrow X$  is strongly measurable if there exist a sequence of simple functions  $\{v_n\}$ , converging pointwise for a.e.  $t$  to  $v$ .

Then, for  $0 < p \leq \infty$ ,

$$L_p(0, T; X) := \{v \mid v \text{ is strongly measurable, and } \|v\|_{L_p(0, T; X)} < \infty\}, \quad (\text{A.0.12})$$

with

$$\|v\|_{L_p(0, T; X)}^p := \int_0^T \|v(t)\|_X^p dt, \quad \text{for } 0 < p < \infty, \quad (\text{A.0.13})$$

and

$$\|v\|_{L_\infty(0, T; X)} := \operatorname{ess\,sup}_{[0, T]} \|v(t)\|_X. \quad (\text{A.0.14})$$

Moreover, the space of continuous functions from  $[0, T]$  into  $X$ , denoted by  $C([0, T]; X)$ , is equipped with the following norm

$$\|v\|_{C([0, T]; X)} = \sup_{[0, T]} \|v(t)\|_X. \quad (\text{A.0.15})$$

We introduce the space of vector distributions  $\mathcal{D}'(0, T; X)$  as the space of continuous linear functionals that maps  $C_0^\infty(0, T) \rightarrow X$ .

## Anisotropic spaces

By the previous section, we note that we have to deal with both time and space directions. We need to be able to have different regularity results in these directions. This is accomplished by the anisotropic spaces introduced in this section.

We let

$$H^s(\mathbb{R}^n) := \left\{ u \in \mathcal{S}'(\mathbb{R}^n) \mid \int_{\mathbb{R}^n} (1 + |\xi|^2)^s \hat{u} \, d\xi \right\}, \quad (\text{A.0.16})$$

and define  $H^s(\Omega)$  to consist of functions in  $H^s(\mathbb{R}^n)$  restricted to  $\Omega$  with norm

$$\|u\|_{H^s(\Omega)} := \inf \left\{ \|\tilde{u}\|_{H^s(\mathbb{R}^n)} \mid \tilde{u} = u \text{ a.e. on } \Omega \right\}. \quad (\text{A.0.17})$$

Similarly,  $H^s(0, T)$  consist of functions in  $H^s(\mathbb{R})$  restricted to  $(0, T)$  with norm

$$\|u\|_{H^s(0, T)} := \inf \left\{ \|\tilde{u}\|_{H^s(\mathbb{R})} \mid \tilde{u} = u \text{ a.e. on } (0, T) \right\}. \quad (\text{A.0.18})$$

As in [LM72a], we let

$$H^{r,s}(Q) := L_2(0, T; H^r(\Omega)) \cap H^s(0, T; L_2(\Omega)) \quad (\text{A.0.19})$$

with the norm

$$\|u\|_{H^{r,s}(Q)}^2 := \int_0^T \|u\|_{H^r(\Omega)}^2 + \|u\|_{H^s(0, T; L_2(\Omega))}^2. \quad (\text{A.0.20})$$

We note that

$$H^{r,s}(\mathbb{R}_t \times \mathbb{R}^n) \text{ is equivalent to } \{v \mid [\langle \xi \rangle^r + \langle \tau \rangle^s] \mathcal{F}_{t,x} v \in L_2(\mathbb{R}_\tau \times \mathbb{R}_\xi)\}. \quad (\text{A.0.21})$$

This can be seen using Fubini's theorem, the Parseval equation, and noting that

$$\begin{aligned} \|w\|_{L_2(\mathbb{R}_t; H^r(\mathbb{R}^n))}^2 &= (2\pi)^{-n} \|\langle \xi \rangle^r \mathcal{F}_x w\|_{L_2(\mathbb{R}_t \times \mathbb{R}_\xi^n)}^2 \\ &= (2\pi)^{-(n+1)} \|\mathcal{F}_t(\langle \xi \rangle^r \mathcal{F}_x w)\|_{L_2(\mathbb{R}_\tau \times \mathbb{R}_\xi^n)}^2 \\ &= (2\pi)^{-(n+1)} \|\langle \xi \rangle^r \mathcal{F}_{t,x} w\|_{L_2(\mathbb{R}_t \times \mathbb{R}_\xi^n)}^2, \end{aligned} \quad (\text{A.0.22})$$

as well as

$$\|w\|_{H^s(\mathbb{R}_t; L_2(\mathbb{R}^n))}^2 = (2\pi)^{-1} \|\langle \tau \rangle^s \mathcal{F}_t w\|_{L_2(\mathbb{R}_\tau \times \mathbb{R}^n)}^2 = (2\pi)^{-(n+1)} \|\langle \tau \rangle^s \mathcal{F}_{t,x} w\|_{L_2(\mathbb{R}_t \times \mathbb{R}_\xi^n)}^2. \quad (\text{A.0.23})$$



# Bibliography

- [Eva08] Lawrence E. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics - AMS, 2008.
- [FCG06] Enrique Fernández-Cara and Sergio Guerrero. Global Carleman inequalities for parabolic systems and application to controllability. *SIAM J. Control Optim.*, 45(4):1395 – 1446, 2006.
- [Gru09] Gerd Grubb. *Distributions and Operators*. Springer, 2009.
- [GS90] Gerd Grubb and Vsevolod Solonnikov. Solution of Parabolic Pseudo-differential Initial-Boundary Value Problems. *Journal of Differential Equations*, 87:256–304, 1990.
- [Had02] Jaques Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52, 1902.
- [Isa98] Victor Isakov. *Inverse problems for partial differential equations*. Springer-Verlag, 1998.
- [Lax02] Peter D. Lax. *Functional Analysis*. Wiley Interscience, 2002.
- [LL67] Jacques-Louis Lions and Robert Lattès. *Méthode de Quasi-Reversibility et Applications*. Dunod, 1967.
- [LM60] Jacques-Louis Lions and Bernard Malgrange. Sur l’unicité rétrograde dans les problèmes mixtes paraboliques. *Mathematica Scandinavica*, 8:227–286, 1960.
- [LM72a] Jacques-Louis Lions and Enrico Magenes. *Non-Homogeneous Boundary Value Problems and Applications*, volume II. Springer-Verlag, 1972.
- [LM72b] Jacques-Louis Lions and Enrico Magenes. *Non-Homogeneous Boundary Value Problems and Applications*, volume I. Springer-Verlag, 1972.
- [Mir61] Willard L. Miranker. A Well Posed Problem For The Backwards Heat Equation. *Proceedings of the American Mathematical Society*, 12(2):243–247, 1961.

- 
- [MZ04] Sorin Micu and Enrique Zuazua. An Introduction to the Controllability of Partial Differential Equations, 2004.
- [Pay75] Lawrence E. Payne. Improperly Posed Problems in Partial Differential Equations, 1975.
- [Paz83] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, 1983.
- [Ped00] Michael Pedersen. *Functional Analysis in Applied Mathematics and Engineering*. Chapman & Hall/CRC, 1. edition, 2000.
- [RS80] Michael Reed and Barry Simon. *Functional Analysis*, volume I. Academic Press, 1980.
- [Rus73] David L. Russel. A Unified Boundary Controllability Theory for Hyperbolic and Parabolic Partial Differential Equations. *Massachusetts Institute of Technology, Studies in applied mathematics*, 52:189–211, 1973.
- [Sho74] Ralph E. Showalter. The Final Value Problem for Evolution Equations. *Journal of Mathematical Analysis and Applications*, 47:563–572, 1974.
- [Tan79] Hiroki Tanabe. *Equations of Evolution*. Pitman, 1979.
- [Tem01] Roger Temam. *Navier-Stokes Equations*. AMS Chelsea Publishing, 2001.
- [Was76] Donald C. Washburn. A Semigroup Theoretic Approach To Modeling Of Boundary Input Problems. *Proceedings of the IFIP working conference*, pages 21–24, 1976.